Situated social cognition:

How contextual factors modulate egocentric and altercentric socio-cognitive processes

Dissertation

zur Erlangung des akademischen Grades

Doctor of Philosophy
(Ph.D.)

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Irene Trilla Gros

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin:
Prof. Dr. Dr. Christian Ulrichs

Gutachter/innen:
1. Prof. Dr. Isabel Dziobek
2. Prof. Dr. Marcel Brass
3. Prof. Dr. Giorgia Silani

Tag der mündlichen Prüfung: 02.09.2021

Acknowledgements

This dissertation would not have been possible without the contributions and support of many people.

First, I would like to thank my first supervisor, Isabel Dziobek, for her guidance and valuable advice over the past years. Since the beginning, she trusted in our projects and encouraged me to find my own path. Her enthusiasm and energy inspired me to move forward, even when challenges did not seem achievable at first. I would like to extend my sincere thanks to Henrik Walter, who co-supervised my doctoral projects with insightful feedback and wise advice. I am also grateful to the Berlin School of Mind and Brain for their financial support and for providing an inspirational environment during my doctorate.

Second, I would like to thank all my colleagues and friends at the Social Cognition Lab. It was a great pleasure to have the opportunity to work alongside Sandra Naumann, Hanna Drimalla, Mareike Bayer, Katharina Bögl, Lena Matyjek, Jan Schneider, Caitlin Duncan, Garret O'Connell, Simón Guendelman, Anne Weigand, Simone Kirst, Silke Lipinski, Renata Wacker, Muyu Lin and Irene Sophia Plank. Beyond their scientific contributions and the many exciting intellectual discussions, they gave me the motivation and support I needed to persist in this journey. I hope they know that, without a doubt, they are the best part of my doctoral experience. I would also like to thank all the students and interns who worked with me and helped completing this project.

Finally, I would like to thank those who encouraged me to pursue my PhD and who supported me along the way. First, Helge Hasselmann, who was there during the first years and without whom I would not have started this journey. Segundo, quiero agradecer a Juan R. Loaiza, por enseñarme a pensar mejor y, sobretodo, por su apoyo durante las etapas más emocionalmente duras del doctorado. També a en Ferran, per la paciència i el recolzament dels últims mesos del treball. I, finalment, vull agrair als meus pares, la Begoña i en Jaume, en Javier i l'Ana, els meus germans, l'Albert i la Dana, i també a la meva àvia Maria i a la meva tieta Pilar per estar sempre amb mi.

Abstract

Social functioning entails flexibly activating and inhibiting different socio-cognitive processes according to the social context. In some circumstances, one's own behaviour and cognition are influenced by the presence of other agents. This is the case of mimicry, the phenomenon whereby a person automatically imitates observed actions and emotional expressions. Other times, it is one's own thoughts and feelings that interfere with how we understand other agents, such as when we project our mental states onto others. This dissertation presents four studies that investigated how altercentric (mimicry) and egocentric (self-projection) processes of social cognition are regulated according to the social context and personal dispositions.

Part I aimed to re-examine the role of social cues previously identified as modulators of mimicry. In line with prior research, Study 1 showed that the tendency to mimic others' happy facial expressions depends on the reward value associated with the observed agent. However, the effects of reward were not in the hypothesised direction, nor could we detect an influence of oxytocin treatment, a hormone that underlies the neurobiology of social adaptation. Study 2 was preregistered to resolve conflicting findings regarding the role of direct and averted gaze cues on the tendency to automatically imitate others' hand actions. We revisited the gaze effects using an automatic imitation task that controlled for the spatial correspondence between the observed and executed hand actions, a confounder that may have affected previous studies. Our data did not reveal a general enhancement of automatic imitation following direct gaze compared to averted gaze. However, we could identify potential contextual and dispositional factors (e.g., autistic traits) that might underlie different mimicry responses to gaze cues.

Part II aimed to test a new framework to study how one's own affective experiences influence our judgments about others' emotions. Combining brief emotion induction blocks with psychophysical measures of emotion perception, Study 3 showed that facial emotional expressions tend to be judged as happier when individuals feel happy than when they feel sad. Emotional egocentric biases were replicated in Study 4, a preregistered experiment that tested the role of perceived similarity on egocentricity. Contrary to our predictions, we did not find stronger egocentric biases when participants judged emotional facial expressions of similar compared to dissimilar others, an effect that had been previously shown for egocentricity in the cognitive domain.

Across all studies, we found evidence supporting the contextual nature of social cognition, even for automatic and low-level processes. However, we could not replicate some of the phenomena reported in the literature. These results highlight the need to systematically re-evaluate the robustness and generalizability of prior findings. Implications of the current work for future research are discussed in the context of the replicability crisis in psychology.

Zusammenfassung

Das soziale Funktionsniveau beinhaltet, je nach persönlichen Zielen und sozialem Kontext, das flexible Aktivieren und Unterdrücken verschiedener sozialer kognitiver Prozesse. Unter bestimmten Umständen werden das eigene Verhalten und die eigene Kognition durch die Anwesenheit und die Handlungen anderer Personen beeinflusst. Dies ist der Fall bei Mimikry, dem Phänomen, bei dem eine Person automatisch beobachtete Handlungen und emotionale Ausdrücke imitiert. In anderen Fällen sind es unsere eigenen Gedanken und Gefühle, die unsere Informationsverarbeitung über andere Personen beeinträchtigen, zum Beispiel dann, wenn wir unsere mentalen Zustände auf andere projizieren. In der vorliegenden Dissertation werden vier Studien vorgestellt, in denen untersucht wurde, wie altrozentrische (Mimikry) und egozentrische (Selbstprojektion) Prozesse der sozialen Kognition in Abhängigkeit vom sozialen Kontext und persönlichen Dispositionen reguliert werden.

Teil I zielte darauf ab, die Rolle sozialer Signale, welche zuvor als Einflussfaktoren der Mimikry-Reaktion identifiziert wurden, erneut zu analysieren. In Übereinstimmung mit früheren Untersuchungen zeigte Studie 1, dass die Tendenz, fröhliche Gesichtsausdrücke anderer nachzuahmen abhängig von dem mit der beobachteten Person assoziierten Belohnungswert ist. Die Auswirkung der Belohnung ging jedoch weder in die vorhergesagte Richtung, noch konnten wir einen Einfluss von Oxytocin, einem Hormon, das der Neurobiologie der sozialen Anpassung zugrunde liegt, finden. In der präregistrierten Studie 2 wurden widersprüchliche Ergebnisse hinsichtlich der Rolle des direkten und abgewandten Blickkontakts auf die Tendenz, die Handbewegung anderer automatisch nachzuahmen, untersucht. Wir überprüften den Einfluss des Blickkontakts mithilfe einer automatischen Nachahmungsaufgabe, die die räumliche Übereinstimmung zwischen beobachteten und ausgeführten Handbewegungen kontrollierte, da dies ein möglicher Störfaktor früherer Studien darstellte. Unsere Daten zeigten keine allgemeine Verbesserung der automatischen Nachahmung nach direktem Blickkontakt im Vergleich zum abgewandten Blick. Wir könnten jedoch potenzielle kontextbezogene und dispositionelle Faktoren (z.B. autistische Eigenschaften) identifizieren, die unterschiedlichen Mimikry-Reaktionen auf den Blickkontakt zugrunde liegen könnten.

Teil II untersuchte, wie die eigenen affektiven Erfahrungen unsere Urteile über die Gefühle anderer beeinflussen. Studie 3 kombinierte kurze Phasen der Emotionsinduktion mit psychophysischen Messungen der Emotionswahrnehmung. Es zeigte sich, dass emotionale Gesichtsausdrücke tendenziell als fröhlicher beurteilt werden, wenn Personen angeben, dass sie sich fröhlich im Vergleich zu traurig fühlen. Emotionale egozentrische Verzerrungen wurden in der präregistrierten Studie 4 wiederholt, in der die Rolle der wahrgenommenen Ähnlichkeit im Kontext der Egozentrik getestet wurde. Entgegen unseren Vorhersagen fanden wir keine stärkeren egozentri-

schen Vorurteile, wenn die Teilnehmenden emotionale Gesichtsausdrücke als ähnlich beurteilten. Dieser Effekt wurde zuvor für die Egozentrik im kognitiven Bereich gezeigt.

In allen Studien fanden wir Hinweise, die die situative Natur der sozialen Kognition für automatische und niedrigstufige Prozesse unterstützen. Allerdings konnten wir einige der in der Literatur berichteten Phänomene nicht replizieren. Diese Ergebnisse unterstreichen die Notwendigkeit, die Robustheit und Generalisierbarkeit früherer Befunde systematisch neu zu bewerten. Im Zusammenhang mit der Replikationskrise in der Psychologie werden Implikationen der aktuellen Arbeit für die zukünftige Forschung diskutiert.

Contents

List of original publications

This dissertation is based on research experiments that have been published or are under revision in peer-reviewed journals.

Study 1:

> Trilla, I., Drimalla, H., Bajbouj, M., & Dziobek, I. (2020). The Influence of Reward on Facial Mimicry: No Evidence for a Significant Effect of Oxytocin. *Frontiers in Behavioral Neuroscience*, *14*. https://doi.org/10.3389/fnbeh.2020.00088

Study 2:

> Trilla, I., Wnendt, H., & Dziobek, I. (2020). Conditional effects of gaze on automatic imitation: The role of autistic traits. *Scientific Reports, 10*(1), 15512. https://doi.org/10.1038/s41598-020-72513-6

Study 3:

> Trilla, I., Weigand, A., & Dziobek, I. (2020). Affective states influence emotion perception: Evidence for emotional egocentricity. *Psychological Research*. https://doi.org/10.1007/s00426-020-01314-3

Study 4:

> Trilla, I., Eiserbeck, F., & Dziobek, I. (2020). *Projecting one's own affective states onto others: No influence of perceived similarity*. PsyArXiv. https://doi.org/10.31234/osf.io/j2ct4

1. General introduction

1.1 Egocentric and altercentric processes of social cognition

Humans have developed a broad range of cognitive functions to navigate the complexities of the social world. Social cognition is the term used to refer to the conjunction of processes that allow us to understand and make inferences about others' intentions, feelings and thoughts, which help to regulate our social behaviour (Happé et al., 2017). While there is no consensus on the taxonomy of socio-cognitive skills, most accounts include components from different levels of social information processing, ranging from emotion and action perception to imitation, empathy, and theory of mind (Chatel-Goldman et al., 2013; Etchepare & Prouteau, 2018; Happé et al., 2017; Lieberman, 2007).

As social beings, our behaviour and cognition are widely influenced by other agents' actions (Kampis & Southgate, 2020). People plan actions, pay attention and react to aspects of their environment motivated not only by their own perspective and goals, but also by those of others around them. For example, we tend to pay more attention to objects when we see other people looking at them (Frischen et al., 2007) and unconsciously imitate others' gestures and facial expressions, even when these interfere with our own action goals (Hess & Fischer, 2013; Heyes, 2011). Moreover, we may feel what another person is feeling despite not having undergone the same experiences (Bird & Viding, 2014; Olszanowski et al., 2020). The phenomenon by which information processing about the self is influenced by the representations of others' states has been referred to as altercentrism. Altercentric social cognition may help align the individual with other group members, thus facilitating interpersonal coordination, communication, group dynamics, and cumulative culture (Kampis & Southgate, 2020).

Contrasting altercentrism, many social cognition theories emphasise the individual as the reference point for accessing the others' minds (Ames, 2004b; Meltzoff, 2007). Even when our attention is ostensibly focused on understanding the thoughts and actions of other agents, our own experiences may intrude (Epley, Keysar, et al., 2004; Silani et al., 2013). The influence of one's own states on the representations related to the others is known as egocentrism. Given the relative similarity within human behaviour, egocentric socio-cognitive processes may help us understand and predict other's actions and cognitions (Epley, Keysar, et al., 2004).

Altercentric and egocentric processes can be seen as complementary processes within social cognition. Given the constant and complex interchange between oneself and others, success-

ful social functioning requires a balance between representing and switching between our own mental states and those of others. While being influenced by others' states and behaviours can aid empathic processes through emotional resonance and mirroring, one also needs to create boundaries between oneself and others to avoid personal distress (Bird & Viding, 2014; Lamm et al., 2016). Similarly, projecting one's own mental states to others can be a useful heuristic to understand other agents. However, it could also lead to biased perceptions and misinterpretations of others' cognition and behaviour unless one can distinguish between oneself and others (Mitchell, 2009; Steinbeis, 2016).

1.2 Situated social cognition

Due to its central role in human behaviour, egocentric and altercentric social cognition has been of interest to many disciplines, including psychology, cognitive science and neuroscience. On many occasions, socio-cognitive abilities are studied as stable, context-unspecific processes, which are assumed to show the same configuration in every situation. However, for successful social functioning, it is not only important that one has a particular socio-cognitive ability, but that this process is activated in a particular situation and at a particular time (Melloni et al., 2014; Smith & Semin, 2007). Imagine, for instance, that a person is helping a friend who is going through a hard time. Showing a certain degree of emotional resonance and empathy is probably beneficial to relate and support their friend. However, if they would be in a competitive situation, such emotional reactivity towards the other's pains could be counter-productive to achieve the personal goals.

In line with a situated notion of social cognition, many socio-cognitive processes are flexibly recruited based on contextual factors (Smith & Semin, 2007). Depending on the environment, interpersonal relationships and goals, social cognition can be oriented towards an egocentric or an altercentric encoding (Kampis & Southgate, 2020). For example, egocentric interferences tend to be stronger when inferring states of in-group than out-group members (Clement & Krueger, 2002; Simpson & Todd, 2017). In visual perspective-taking tasks, individuals are faster at judging the self-perspective when one's own and the other's perspectives are in conflict (Samson et al., 2010). In contrast, the other's perspective is judged faster when there is no conflict.

Recognising the importance of context as an inherent explanatory variable is not only important to understand the mechanisms underlying social cognition but may have clinical implications as well. For example, impairments in social functioning in autism spectrum conditions were initially related to general deficits in empathy and the mirror neuron system (Bernier et al., 2007; Dziobek et al., 2008; Iacoboni & Dapretto, 2006; McIntosh et al., 2006). However, more recent studies have shown that individuals with ASC can show empathic and mirroring responses (Hamilton, 2013; Hamilton et al., 2007; Hudry & Slaughter, 2009; Oberman et al., 2008). Rather than an inability to exert empathic behaviour, individuals with

ASC may adjust their social behaviour differently to the interpersonal context compared to neurotypical individuals (Chevallier et al., 2012; Wang & Hamilton, 2012). Understanding socio-cognitive processes as malleable and sensitive to current social situations could help better characterise deficits in social functioning.

In this dissertation, I will present a series of studies that examined how contextual factors influence egocentric and altercentric processes of social cognition. Study 1 and Study 2 analysed the impact of social cues on automatically mimicking others' facial expressions and actions (altercentric process). Study 3 and Study 4 focused on egocentric processes of affective inferences and how they are modulated by the perceived similarity with the other.

1.3 Part I – Altercentrism: Mimicry

1.3.1 What is mimicry?

Mimicry is the spontaneous imitation of others' actions, postures and facial expressions. Contrarily to intentional forms of imitation, in which an individual deliberately replicates the actions of another person, mimicry is an automatic process that occurs without conscious awareness. Individuals do not notice when they mimic others nor when others mimic them (Chartrand & Lakin, 2013; Chartrand & van Baaren, 2009; Duffy & Chartrand, 2015).

Researchers have distinguished different forms of mimicry, depending on the type of imitated action (Duffy & Chartrand, 2015). Hess and Fischer (2014) differentiate between behavioural and emotional mimicry. The first refers to postures, gestures or facial expressions with a neutral context (e.g., foot-tapping, face touching) and the latter to nonverbal behaviours that express an emotional state (e.g., emotional facial expressions). Other research has distinguished between behavioural mimicry of social vs non-social action (Cracco, Genschow, et al., 2018). The former refers to gestures that do not inherently express an emotion, but are socially meaningful, such as thumbs up or raising the middle finger (Bourgeois & Hess, 2008; Cracco, Genschow, et al., 2018). These distinctions are important when assessing the roles and modulators of mimicry, as different functions can be expected for mimicry of communicative gestures or expressions, compared to actions that do not convey any particular social meaning.

Despite being slightly different phenomena, theoretical accounts on the different manifestations of mimicry assume common underlying processes. It is generally agreed that mimicry is based on a direct mapping of observed and executed actions (Chartrand & Bargh, 1999; Heyes, 2011). The human mirror neuron system has been proposed as the neural basis supporting mimicry behaviour, as it is activated both when executing an action and when observing the same action being performed by another person (Brass & Heyes, 2005; Campbell & Cunnington, 2017; Catmur et al., 2009; Heyes, 2011; Likowski et al., 2012).

1.3.2 What is the function of mimicry?

Mimicry is thought to act as a social regulator by promoting affiliation and reinforcing social bonds (Hess & Fischer, 2013; Lakin & Chartrand, 2003; Wang & Hamilton, 2012). Individuals mimic more those whom they like and like more those who mimic them (Duffy & Chartrand, 2015; McIntosh, 2006; Neufeld & Chakrabarti, 2016; Stel et al., 2009), a bidirectional link that strengthens the interpersonal relationship (van der Schalk, Fischer, et al., 2011). Mimicry can also be used to ameliorate unfavourable social situations. For example, studies have shown that people exhibit higher mimicry responses to interaction partners that had previously excluded them, which has been interpreted as a strategy to recover from the experience of being excluded and re-establish the social bond (Lakin et al., 2008).

In addition to its affiliative role, mimicry seems to facilitate emotion recognition and action understanding (Drimalla et al., 2019). According to embodied cognition theories, spontaneous mimicry reflects the internal simulation of the perceived emotion, which facilitates the understanding of the other's affective state (Niedenthal, 2007). This claim is supported by studies showing that blocking the possibility to mimic the observed emotional expression impairs its recognition (Niedenthal et al., 2010; Oberman et al., 2007; Stel & van Knippenberg, 2008).

Finally, by linking perception and action, mimicry is likely to play a role in social learning, coordination and synchrony between individuals (Chartrand & Bargh, 1999; Heyes, 2011). According to the implicit socialisation account, mimicry may support the learning of implicit group- and environment-specific skills and actions (Kavanagh & Winkielman, 2016). Automatically mimicking others, individuals learn how the social group reacts and implicitly assimilate their gestures, postures, expressions, accents, and mannerisms.

1.3.3 How is mimicry studied experimentally?

Different approaches have been used to study mimicry in the lab. Social psychologists typically measure the frequency of mimicry of foot-tapping or face touching when participants are interacting with a confederate (e.g., Chartrand & Bargh, 1999; Lakin & Chartrand, 2003). Such naturalistic studies have high ecological validity but are limited in experimentally controlling contextual factors, such as the emotional displays or eye contact of the interaction partner.

To obtain higher experimental control, cognitive psychology developed a laboratory model of mimicry, usually referred to as automatic imitation (Heyes, 2011). Automatic imitation is measured with stimulus-response compatibility tasks in which the participant is asked to perform a motor movement while simultaneously observing a congruent or incongruent action (Brass et al., 2000). Mimicry is indicated if the congruent observed action facilitated the participant's motor movement, and/or if the incongruent action interfered with their movement. Automatic imitation paradigms offer high experimental control, as the stimuli's

characteristics, timings, and set of actions to be observed and executed can be manipulated. Moreover, the outcome measures (i.e. reaction times, accuracy) are usually more easily quantifiable and sensitive than more naturalistic measurements of behavioural mimicry.

A third form of mimicry that has capture researchers' attention is (emotional) facial mimicry, which refers to the spontaneous matching of the other's emotional facial expressions. Facial mimicry is distinguished from other affective processes that lead to congruent facial reactions, such as emotional contagion or affective empathy, in that the mimicked facial expressions reflect the sharing of the emotional displays, rather than a response to the other's emotional state (Hess & Blairy, 2001; Hess & Fischer, 2014). Facial mimicry responses are subtle and difficult to observe visually. However, they can be sensitively indexed by electromyographically (EMG)-recorded activity in facial muscles (Hess, Arslan, et al., 2017; van Boxtel, 2010), or by scoring facial movements from video recordings according to the Facial Action Coding System (Sato & Yoshikawa, 2007). More recently, automated facial recognition software has also been used to track facial mimicry responses (Drimalla et al., 2020).

This dissertation will analyse the literature on behavioural mimicry, automatic imitation and facial mimicry jointly, as these phenomena are thought to fulfil comparable functional roles. However, it is still unclear how the different forms of mimicry relate to each other (Genschow et al., 2017). I acknowledge that certain aspects may differ, especially concerning their social modulation. In these cases, differences will be highlighted throughout the topic's discussion.

1.3.4 Social modulation of mimicry

Early theories understood mimicry as a stimulus-driven, reflex-like phenomenon, where the mere observation of an action would automatically trigger a matching nonverbal display in the observer (Chartrand & Bargh, 1999). This notion has been readjusted to the observation that mimicry is not initiated in all circumstances in which a person perceives someone else's actions. Even if it happens without conscious awareness, people seem to preferentially mimic others' actions and emotional displays when there is a minimal potential of engagement and inhibit mimicry when it would not lead to any social advantage (Fischer & Hess, 2017; Genschow & Schindler, 2016). Recent accounts have thus moved towards understanding mimicry as a motivated, goal-directed behaviour, the occurrence of which is conditioned to particular social contexts (Chartrand & Lakin, 2013; Fischer & Hess, 2017; Wang & Hamilton, 2012).

If mimicry is a strategy for social advantage, factors related to both the perceiver, the target, their relationship, and their social goals will likely play a role in determining whether an action, gesture or emotional expression will be mimicked or not (Seibt et al., 2015). In this dissertation, I revisited two aspects of the social context that have been shown to influence

the degree of mimicry: the reward value of the observed agent (Study 1) and their gaze cues (Study 2).

1.3.4.1 Implicit attitudes & reward

Because being mimicked tends to be implicitly perceived as an affiliative signal, motivational theories predict that people imitate the actions and emotional expressions of individuals they hold a positive attitude towards, presumably because this will lead to a higher social benefit (Hess & Fischer, 2013; Wang & Hamilton, 2012). In contrast, mimicking someone with whom one has a negative relationship or does not want to bond could be seen as socially inappropriate and lead to negative consequences. In line with this idea, research has shown that people tend to preferentially mimic agents they like and inhibit mimicry when interacting with a disliked person (Blocker & McIntosh, 2016; Likowski et al., 2008; McIntosh, 2006; Stel et al., 2009). Similarly, mimicry tends to be up-regulated when observing actions or specific emotional reactions of in-group members compared to out-group members (Bourgeois & Hess, 2008; Guéguen & Martin, 2009; van der Schalk, Fischer, et al., 2011; Yabar et al., 2006).

The influence of implicit attitudes has also been shown in experiments in which the reward value associated with the observed agent was manipulated. Sims et al. (2012) associated different faces to losing and winning monetary rewards using an implicit conditioning paradigm. Afterwards, participants observed emotional facial expressions displayed by the conditioned agents, while mimicry responses were recorded using electromyography. Congruent with motivational theories of mimicry, happy expressions by agents associated with rewards were more strongly mimicked than those of agents paired with losing money (Sims et al., 2012). This finding has been replicated in subsequent mimicry studies that used similar versions of the reward conditioning paradigm (Haffey et al., 2013; Korb et al., 2019).

Arguably, these reward manipulations altered the implicit attitudes towards the stimuli, with faces linked to higher rewards being more liked (Korb et al., 2019). Rewards have also been shown to enhance mimicry by modulating the motivation to engage with the interaction partner. For example, sad expressions were mimicked only if participants obtained monetary rewards for accurately identifying the emotional expressions displayed (Hess, Blaison, et al., 2017). When participants did not expect any reward, viewing sad faces did not elicit a matching facial reaction. Altogether, evidence suggests that the reward value ascribed to an interactant determines how much we will mimic their emotional expressions and gestures. This is congruent with the idea that mimicry is driven by the motivational drive to interact with others (Lakin & Chartrand, 2003; Wang & Hamilton, 2012).

Study 1 – Research aim (1)

Study 1 aimed to re-examine the influence of reward on facial mimicry using a modified version of the task developed by Sims et al. (2012). Sims et al. (2012) used a classical conditioning paradigm to implicitly associate different face identities with monetary rewards and punishments. This task proved to be effective in manipulating the value of the stimuli (Korb et al., 2019). However, it had the limitation that the learning of the pairings between the face stimuli and monetary outcomes was incidental to the participant's task and could not be tracked.

To fill this gap, Study 1 implemented a reward learning task in which the monetary rewards paired with the face stimuli were contingent on the participant's task performance. This allowed us to verify that the reward associations were assimilated before measuring facial mimicry reactions when observing happy and angry expressions displayed by the conditioned face identities. To avoid confounding the effects of punishments (losing money) vs rewards (gaining money), our task compared only variations in facial mimicry to high vs low reward probabilities of winning money.

1.3.4.2 Gaze

In addition to the role of interpersonal attitudes and reward, mimicry has been shown to depend on affiliative signals displayed by the interaction partner. Emotional expressions and gaze cues are a rich source of non-verbal information that help decode others' intentions and internal states during social interactions (Hamilton, 2016). Establishing direct gaze is often interpreted as a sign of social interest and intention to engage with the perceiver (Wirth et al., 2010). Gaze cues also play an important role in social referencing, as they shift the attentional focus toward the gaze direction and elicit joint attention (Frischen et al., 2007). Given its socially communicative function, it is not surprising that gaze direction has been studied as a factor determining mimicry.

As a cue of social approachability, prior research had hypothesised that individuals would preferentially mimic agents that established direct eye contact, compared to agents who averted the gaze away from the interaction. As predicted, a series of studies observed stronger mimicry when observing hand actions and smiles of individuals with direct gaze compared to averted gaze (Forbes et al., 2016; Schrammel et al., 2009; Wang, Newport, et al., 2011; Wang & Hamilton, 2014). While evidence of an enhancement of mimicry following direct gaze seemed consistent across reports, at least two studies could not replicate these observations (Farmer et al., 2021; Marsh et al., 2016).

Crucially, the study by Marsh et al. (2016) used a task that disentangled the influence of gaze on imitative and spatial compatibility, two dissociable effects that are confounded in certain types of automatic imitation paradigms (Boyer et al., 2012; Catmur & Heyes, 2011; Cracco,

Bardi, et al., 2018). With this set-up, Marsh et al. (2016) did not observe gaze-related modulations of automatic imitation. Instead, the target's gaze direction and ethnicity specifically influenced spatial compatibility. Given that the automatic imitation task used in the original studies on the influence of gaze could not assess the independent effects of imitative vs spatial compatibility, findings by Marsh et al. (2016) opened the possibility that the assumed role of direct and averted gaze cues was not directly influencing mimicry but other processes of response inhibition.

In addition, most automatic imitation studies on gaze included only stimuli depicting a single facial identity in their paradigm (e.g., Forbes et al., 2017; Wang, Newport, et al., 2011; Wang & Hamilton, 2014). Because factors such as gender, ethnicity and emotional facial expressions can influence how gaze cues are interpreted (Adams Jr. & Kleck, 2005; Ohlsen et al., 2013; Weisbuch et al., 2017), particular attributes of the stimuli used could have contributed to the effects of direct and averted gaze.

Study 2 – Research aim (1)

Study 2 re-examined the effects of direct and averted gaze on automatic imitation with the aim to resolve the inconsistent findings regarding the role of gaze on mimicry. The experiment was designed to control for experimental factors that could have confounded the gaze effects in earlier studies. First, as in Marsh et al. (2016), we used paradigm that disentangled the measurement of automatic imitation from spatial compatibility. Second, we developed a stimulus set that included multiple facial identities and which allowed to control for factors such as the targets' gender, ethnicity, and the direction of the head movements that determined the gaze cues.

1.3.5 Dispositional factors influencing mimicry

In addition to contextual determinants, research has identified that the tendency to mimic others depends on individual predispositions. A recent meta-analysis of facial mimicry studies revealed significant associations between empathy and the tendency to imitate others' emotional expressions (Holland et al., 2021). In particular, people with higher empathic dispositions or who report sharing the others' emotional state to a greater degree tend to show stronger facial mimicry (Dimberg et al., 2011; Drimalla et al., 2019; Sonnby–Borgström, 2002). In contrast, non-emotional forms of mimicry, such as automatic imitation, seem more invariant to socio-emotional traits. A meta-analysis with a combined sample of almost 1000 participants did not find any significant associations between automatic imitation and trait empathy (Cracco, Bardi, et al., 2018). Moreover, a high-powered study also failed to find associations between automatic imitation and a series of personality traits, including agreeableness, extraversion, autistic traits, schizotypal traits, narcissistic personality, alexithymia and dispositional empathy (Darda et al., 2020). The fact that traits related to social-emotional

processing seem to be more determinant for facial mimicry than for other non-emotional forms of imitative behaviour is not entirely surprising, given that facial expressions are a communicative and affective signal.

Beyond individual differences in the general tendency to mimic others, people may differ in how they regulate their mimicry behaviour to the current context. If we conceive mimicry as a motivated and strategic behaviour, personal characteristics related to how one processes and responds to contextual social cues could be related to individual differences in the social modulation of mimicry. Study 1 and Study 2 aimed to examine the role of two personal factors proposed to determine the contextual regulation of mimicry: oxytocin (Study 1) and autistic traits (Study 2).

1.3.5.1 Oxytocin

Oxytocin is an evolutionary ancient hormone and neuromodulator that has been involved in several physiological and psychological functions that regulate social and non-social human behaviour and cognition (Feldman et al., 2016; Harari-Dahan & Bernstein, 2014). It is primarily produced in the hypothalamus and secreted both within the brain and into the circulatory system (Meyer-Lindenberg et al., 2011). As an allostatic neuropeptide, it is thought to support behavioural strategies to adapt to changing environmental conditions (Quintana & Guastella, 2020). To improve the capability for social adaptation, oxytocin may increase the salience of socially relevant stimuli (Shamay-Tsoory & Abu-Akel, 2016), promote social motivation (Bethlehem et al., 2014) and regulate affective states (Ma et al., 2016).

Intranasal oxytocin administrations have become one of the most used methods to study the role of this neuropeptide in human behaviour and cognition. Because of its size, oxytocin cannot cross the blood-brain-barrier and reach the brain in appreciable amounts when administered peripherally or into the circulatory stream (Quintana et al., 2015). Taking advantage of the direct pathways between the nasal cavity and the central nervous system, nasal spray delivery is an easy and non-invasive route to manipulate central oxytocin levels (Quintana et al., 2018).

Owing to its involvement in social behaviour and cognition, oxytocin has captured interest in mimicry research (Kraaijenvanger et al., 2017). An early study found lower inhibition of automatic imitation following intranasal oxytocin treatment (De Coster et al., 2014). However, this result could not be replicated in a more recent experiment that used a similar automatic imitation task (Tomova et al., 2019). Oxytocin effects on emotional mimicry are also mixed. Korb et al. (2016) reported marginal oxytocin-driven increases in facial mimicry when judging emotional expressions of infant faces, but not of adults. In contrast, Pavarini et al. (2019) found up-regulation of mimicry of adult sadness and happiness following intranasal oxytocin, although the latter only in individuals who showed low positive

expressivity. Furthermore, no significant effects of oxytocin were detected for mimicry of anger.

Altogether, findings from existing studies show tentative evidence for a role of oxytocin on mimicry. However, the pattern of results is complex, and effects seem highly dependent on contextual factors, such as the observed emotional expression and the age of the interactant. Considering that current theories understand oxytocin as a regulator of (social) behaviour (Ma et al., 2016; Quintana & Guastella, 2020), intranasal oxytocin may work by enhancing or inhibiting mimicry according to the requirements of the social context.

Study 1 – Research aim (2)

The second aim of Study 1 was to test the role of oxytocin on facial mimicry. Based on the idea that oxytocin is involved in the regulation of social behaviour, we predicted that intranasal oxytocin would specifically influence the reward-related modulation of facial mimicry, rather than generally increasing or decreasing mimicry behaviour. In particular, we hypothesised stronger imitation of emotional expressions by faces paired with high rewards vs low rewards following oxytocin treatment compared to placebo.

1.3.5.2 Autistic traits

Autism spectrum conditions (ASC) are etiologically and clinically heterogeneous developmental disorders characterised by impairments in social interaction, communication, restricted interests, and repetitive behaviours (American Psychiatric Association, 2013). Amongst the alterations in social functioning, research has identified atypical patterns of mimicry behaviour (Forbes et al., 2016; Oberman et al., 2009). Initially, mimicry impairments were thought to be linked to a mirror neuron system dysfunction (McIntosh et al., 2006; Oberman & Ramachandran, 2007). However, the "broken mirror system hypothesis" was challenged by subsequent studies, which could not detect consistent deficits in the ability to spontaneously mimic others' facial expressions or gestures in ASC (Cracco, Bardi, et al., 2018; Hamilton et al., 2007; Press et al., 2010; Schulte-Rüther et al., 2017).

Even though individuals with ASC may not exhibit fundamental mimicry alterations, some studies have reported atypical mimicry regulation to the social context. For example, individuals with ASC do not seem to regulate their imitative behaviour according to the gaze cues of the observed agent (Forbes et al., 2017; Vivanti & Dissanayake, 2014). Similarly, mimicry behaviour in ASC was not enhanced by prosocial priming compared to non-social priming, an effect that was found in the control sample (Cook & Bird, 2012). Autistic traits have also been shown to determine how contextual factors influence facial mimicry in the general population. For example, the reward effect observed in Sims et al. (2012) was stronger for individuals with low autistic traits. In contrast, mimicry reactions by individuals with higher autistic traits were independent of the reward associated with the mimicked agent.

Haffey et al. (2013) also observed increased automatic imitation for hands associated with high rewards vs low rewards, but only in individuals with low autistic traits or high empathy.

Taken together, current literature suggests a role of autistic traits in determining the extent to which mimicry behaviour is flexibly adjusted to the social context. These findings are congruent with the social top-down response modulation theory, which proposes that mimicry impairments in ASC are related to reduced adaptability to the social environment (Wang & Hamilton, 2012).

Study 2 – Research aim (2)

The second aim of Study 2 was to examine whether mimicry patterns to direct and averted gaze cues are conditional on autistic traits in the general population. Although there is an ongoing debate on whether ASC is best viewed as an extreme of a neurodevelopmental continuum or as a distinct category, studies have shown that autistic traits across the general population are etiologically linked to autistic traits in individuals with ASC (Lundström, 2012; Robinson et al., 2011). Given that quantitative measures of autistic traits seem to assess the same latent constructs in ASC and non-clinical samples (Murray et al., 2014), studying associations with autistic traits in subclinical populations could help identify relevant phenomena for ASC.

1.4 Part II – Egocentrism: Self-projection

1.4.1 What is self-projection?

Much of our social life is based on our ability to infer the mental states of others. Due to a lack of direct access to other people's minds, we need to intuit what others think, want or feel by interpreting their actions and emotional expressions. However, in many situations, others' overt behaviour may be ambiguous, or we may just not have enough evidence about their state to base our social inferences on. In such cases, we will need to rely on more indirect sources of information to understand others' mental states.

One common strategy is to use the self as the reference point to access others' minds. For instance, we may predict if someone will like a film based on how much we liked it; or decide to offer someone a warm blanket because we are ourselves cold. People tend to assume that other people think, feel and behave as they do (Epley, Keysar, et al., 2004; Mitchell, 2009). The salience of self-knowledge often causes individuals to overestimate the prevalence of their own preferences, behaviours and traits, as shown by the widely replicated false consensus effect (Klein et al., 2018; Ross et al., 1977). The tendency to use self-knowledge to predict and understand others' mental states is referred to as self-projection or egocentrism (Epley, Keysar, et al., 2004; Mitchell, 2009; Van Boven & Loewenstein, 2003).

Self-projection has been shown to occur in various kinds of situations and inferred content: from making inferences about "stable" characteristics, such as personality traits, preferences or attitudes (Critcher & Dunning, 2009; Davis et al., 1996; Ready et al., 2000; Tamir & Mitchell, 2013); to inferences about "situated" mental states, such as beliefs, intentions and emotional reactions elicited under particular circumstances (Ross et al., 1977; Van Boven & Loewenstein, 2003; Yik et al., 2019). Especially when inferring "situated" states, self-projection may be an inherent part of a perspective-taking process (Epley, Keysar, et al., 2004; Mitchell, 2009). To assess how another person is thinking or feeling in a specific context, we may first imagine ourselves in that situation and assess how we would react. We then infer the other's state by projecting our assumed responses.

Given the relative similarities between all human minds, basing social inferences on our own subjective experiences can be an efficient heuristic, especially when other sources of information are limited. However, self-knowledge should represent only the anchor or starting point of the inferential process. An adjustment mechanism is needed to account for dissimilarities between oneself and the other person (Epley, Keysar, et al., 2004; Mitchell, 2009; Tamir & Mitchell, 2013). For example, even though we liked a film, we may end up not recommending it to a friend if we know that the other person is not a fan of the film's genre. To make accurate inferences, an adjustment from the initial anchor is necessary to correct away from the self-knowledge based on the available information about the other and their context (Tamir & Mitchell, 2013). A failure in this correction process could lead to egocentric biases, which are inaccuracies in social inferences due to an over-attribution of one's own mental states. Egocentric judgments will be particularly evident when we make inferences about people who are in a different state than ours (Silani et al., 2013). The greater the difference between ourselves and the other person, the greater the adjustment that is needed.

1.4.2 How is self-projection studied experimentally?

Broadly, investigations on self-projection have used two different approaches. On the one hand, social psychology research has mostly relied on paradigms in which participants make judgments about their traits, attitudes or behaviours along a particular dimension and then estimate the position of another individual on that dimension (e.g., Critcher & Dunning, 2009; Ready et al., 2000). An overestimation of the occurrence of their own choice is taken as an indication of self-projection. However, positive correlations between one's own and inferred traits or attitudes of others may not always indicate that the self is used as a proxy for social inferences (Bazinger & Kühberger, 2012). Imagine, for example, that we predict that a friend will enjoy eating chocolate. If we also like chocolate, the correlation between our own and the predicted preference would be high. However, our inference could have actually been based on a theory that most people like chocolate, a general belief that, in this case, happens to be congruent with our own preference. To overcome this issue, experiments that use correlations between self and other attributes to assess self-projection need to in-

clude adequate control conditions, such as comparing the covariation against an alternative inferential strategy (Bazinger & Kühberger, 2012; e.g., Ames, 2004b; Tamir & Mitchell, 2013).

A different approach to investigate self-projection is to use tasks in which participants infer "situated states" of others while being simultaneously exposed to matching or distinct conditions. False belief paradigms and perspective-taking tasks are examples of this second approach. Common false belief paradigms test whether one's own privileged knowledge of the location of an object biases the prediction of a protagonist's false belief about the object's location (e.g., Coburn et al., 2015; Ryskin & Brown-Schmidt, 2014). In this task, inferences about the protagonist's behaviour tend to be biased because the protagonist's beliefs differ from the participant's own knowledge about the situation.

In perspective-taking tasks, participants are usually confronted with a situation that elicits congruent or incongruent perceptions compared to the situation the target person is exposed to. For example, in visual perspective-taking tasks, participants are asked to make judgments about what another person is seeing (Epley, Morewedge, et al., 2004; Samson et al., 2010). Crucially, some of the stimuli are occluded from the target's point of view and are only visible to the participant. In order to unbiasedly infer the other's perceptions, the participant needs to disregard their own view. An intrusion of our own visual input when judging the other's visual experience is taken as evidence of self-projection and egocentricity (Apperly et al., 2010; Bukowski & Samson, 2017; Keysar et al., 2003; Samson et al., 2010).

1.4.2.1 Emotional egocentricity paradigms

So far, research on self-projection has mainly focused on investigating egocentric biases during inferences of cognitive states, understood here as beliefs, traits or attitudes. Comparably, only a few studies have investigated whether individuals also project their own affective states when predicting how another person is feeling. Most current emotional egocentricity paradigms follow the perspective-taking approach. Participants are asked to predict how another person is feeling in a particular situation, while being in an affective state congruent or incongruent to that of the target (e.g., Silani et al., 2013; Van Boven & Loewenstein, 2003). Evidence of self-projection is indicated if the participant's current state biases the evaluations of the other agent's state.

Emotional egocentric biases have been consistently detected with paradigms in which participants are induced positive or negative feelings via visuotactile (Silani et al., 2013), visual-gustatory (Hoffmann et al., 2015) or audio-visual stimulation (von Mohr et al., 2019), or monetary reward and punishment (Steinbeis & Singer, 2014). Participants are asked to rate the experience of another person who is supposedly undergoing an affectively congruent or incongruent stimulation simultaneously. In line with the idea of self-projection, judgements about the other's affective state have been shown to be biased towards self-experiences, an

effect that tends to be stronger in incongruent conditions (Hoffmann et al., 2015; Silani et al., 2013; Steinbeis & Singer, 2014; Van Boven & Loewenstein, 2003; von Mohr et al., 2019).

Notably, in current affective self-projection paradigms, emotion attributions are made based on contextual information about the situation the other person was exposed to. Participants cannot see the target's emotional responses to the stimulation, so predictions have to be based solely on the provided information about the described event. However, in daily life situations, we can often rely on the other person's overt behavioural reactions and emotional expressions to understand what they are feeling, an ability referred to as emotion perception. We may infer that a person is happy not only because we know they received very positive news, but also because they are smiling and speaking cheerfully. Compared to affective perspective-taking and theory of mind, emotion perception is thought of as a more automatic and implicit form of emotion processing (Etchepare & Prouteau, 2018; Mier et al., 2010; Tracy & Robins, 2008). Despite the relevance of emotion perception for affective inferences, most literature on self-projection has only focused on investigating emotional perspective-taking processes.

Study 3 – Research aim

Study 3 aimed to develop a paradigm to investigate self-projections during emotion perception, a lower-level process of emotion attribution than the type of affective inferences studied with prior emotional egocentricity paradigms. In particular, we examined whether evaluations of others' emotional facial expressions are biased towards the observers' own affective states. If self-projection occurs during emotion perception, individuals would tend to judge facial expressions as happier when they are feeling happy compared to when they are sad.

1.4.3 Social modulation of self-projection

While self-projection is a common mechanism for social inference, people do not always default to using the self as a proxy. According to the similarity contingency model (Ames, 2004b), individuals may deploy or withhold self-projection depending on the perceived similarity with the other agent. When a person assumes high similarity with the target, they tend to ascribe their own attitudes, traits and beliefs to the other. If the perceived similarity with the other is low, anchoring on the self would require a significant correction process to move away from the egocentric attributions. In those cases, individuals tend to rely on other sources of information as a starting point for social inferences.

The similarity-contingency model is supported by studies showing that people tend to project the self-knowledge more when making inferences about similar individuals or groups but recruit less or no self-projection when attributing mental states to dissimilar agents (Ames, 2004b, 2004a; Davis, 2017). The similarity effects do not revolve around the actual

degree of similarity, but on initial subjective evaluations of the similarities with the target, which could even be founded on factors of little validity or relevance to the inference at stake (Ames, 2004a). For example, effects have been found with lab manipulations in which participants just learnt that they shared a few specific attributes with a target (Ames, 2004a), or in which participants were primed to pay attention to the similarities or differences with various targets groups (Ames, 2004b). Subjective perceptions of similarity may also explain the effects of social categorisation observed in social projection paradigms, which found consistent evidence for egocentric inferences about in-group members, but not out-group members (Clement & Krueger, 2002; Robbins & Krueger, 2005).

Self-knowledge does not seem to be recruited as an "anti-anchor" when making attributions about dissimilar agents (Clement & Krueger, 2002; Robbins & Krueger, 2005). That is, we do not predict opposing mental states to dissimilar others (e.g., "As I do not like reading, a dissimilar person will probably like reading"). Instead, people use alternative inferential strategies, such as relying on implicit beliefs or stereotypes about a particular group as a starting point. For instance, if we sense that our mother-in-law's preferred literature genres differ from ours, we may employ stereotypes about the types of books people her age typically read to decide which book to gift her. Alternatively, we could anchor on a well-known person in our lives that we think is more similar to the target than ourselves (Clement & Krueger, 2002)(Willard & Markman, 2017). That is, instead of basing our predictions on stereotypical information, we may choose the book based on what our partner would like, as we know they and their mother have a similar taste in books. Notably, perceived similarity does not necessarily function as a dichotomous on-off switch between inferential strategies. Varying degrees of self-projection and stereotyping can be recruited when making attributions for both similar and dissimilar agents (Ames, 2004b).

So far, the effect of similarity on social projection has been mainly investigated for inferences of others' traits, attitudes, and preferences. By contrast, studies assessing the role of similarity on affective self-projection are scarce. The only evidence that perceived similarity may condition the degree to which we project our own affective states to others is provided by O'Brien and Ellsworth (2012). The authors detected egocentric biases when making judgments about the feelings of thirst and coldness of a similar target, but not when judging dissimilar targets. Specifically, participants who ate salty snacks without water or who were outside during winter overestimated the feelings of thirst and coldness of a protagonist of a story, but only if the target shared the same political views as the participant.

Study 4 – Research aim

Study 4 aimed to test whether perceived similarity determines the degree of affective self-projection during emotion perception. The degree to which the own affective states biased emotion judgements of similar and dissimilar agents was measured using the emotional egocentricity paradigm developed in Study 3. A sense of (dis)similarity with the evaluated agents was created by assigning the same or a different arbitrary attribute to the participant and the agents during the experimental session.

1.5 Summary of research aims and hypotheses

This dissertation aimed to investigate how socio-cognitive processes are regulated according to contextual factors and personal dispositions. To achieve a more comprehensive perspective on the situated nature of social cognition, this work investigated two complementary socio-cognitive processes. On the one hand, Study 1 and Study 2 (part I) focused on the tendency to mimic others' emotional expressions and hand actions (*altercentrism*) and how its occurrence depends on the social context. On the other hand, Study 3 and Study 4 (part II) examined whether one's own affective states interfere with our evaluations about other's emotional expressions (*egocentrism*).

1.5.1 Part I – Social modulation of mimicry

The research objective of part I was twofold. First, we aimed to conceptually replicate the effects of social cues previously identified as modulators of mimicry, namely, reward (Study 1) and gaze (Study 2). Second, we investigated personal factors that could be related to individual differences in the contextual regulation of mimicry, including oxytocin levels (Study 1) and autistic traits (Study 2).

Study 1: Facial mimicry, reward & oxytocin

Study 1 examined the role of oxytocin on the reward-driven modulation of facial mimicry. After intranasal administrations of oxytocin or placebo, participants watched happy and angry facial expressions displayed by agents previously paired with high and low reward probabilities. The participants' tendency to mimic emotional expressions was assessed using facial EMG. Based on previous research (Korb et al., 2019; Sims et al., 2012), we hypothesised that facial mimicry would be stronger when observing emotional expressions of agents associated with high vs low reward probabilities. In addition, we predicted that the oxytocin treatment would enhance the influence of reward on facial mimicry compared to placebo.

Study 2: Automatic imitation, gaze & autistic traits

Study 2 investigated the impact of gaze cues on automatic imitation. To measure automatic imitation, participants were asked to perform finger movements while simultaneously observing another agent executing an action that could be imitatively and/or spatially congruent or incongruent. In some trials, the observed agent had previously directed the gaze towards the participant. In other trials, the agent averted the gaze away from the participant. Based on previous literature, we expected stronger automatic imitation following direct gaze than averted gaze (Wang, Newport, et al., 2011; Wang & Hamilton, 2014). Alternatively, gaze could selectively affect spatial compatibility, such as found in Marsh et al. (2016). In this case, direct gaze was expected to enhance spatial compatibility effects compared to averted gaze. Additionally, we predicted that the influence of gaze on automatic imitation would be conditional to individual differences in autistic traits. In particular, we hypothesised weaker gaze effects with increasing levels of autistic traits.

The hypotheses, methods and statistical analyses of Study 2 were preregistered (https://osf.io/84wqe).

1.5.2 Part II – Emotional egocentricity

Studies included in part II aimed to test a new framework to investigate affective self-projection. Complementing existing emotional egocentricity paradigms, which assessed self-projection during processes of affective perspective-taking (e.g., Silani et al., 2013; von Mohr et al., 2019), Study 3 and Study 4 examined the occurrence of egocentric biases in emotion perception. Moreover, we assessed whether the extent of emotional egocentricity would vary according to individual differences in socio-emotional traits (Study 3) and the perceived similarity with the evaluated agents (Study 4).

Study 3: Emotional egocentricity & socio-emotional traits

Study 3 tested whether one's own affective states egocentrically bias evaluations of others' emotional facial expressions. In a within-subject design, participants were induced transient positive, negative and neutral affective states. After each emotion induction, a psychophysical procedure was used to measure the extent to which their affective state biased judgments about ambiguous emotional expressions. We hypothesised that participants would judge the facial expressions as happier when being in a positive state than when feeling sad. In addition, we explored associations between the degree of emotional egocentric biases and dispositional empathy and autistic traits.

Study 4: Emotional egocentricity & similarity

Study 4 was planned as an extension of Study 3 to investigate whether the degree of emotional egocentricity was conditional on the perceived similarity with the other. Similarity was experimentally manipulated by assigning the same or a different bogus cognitive style

to the participant and a series of targets. Using the emotional egocentricity paradigm developed in Study 3, we measured the extent to which individuals were biased by their own affective states when judging emotional facial expressions of similar and dissimilar targets. Based on the findings in Study 3, we hypothesised a higher tendency to judge facial expressions as happy following positive vs negative affect induction. Moreover, emotional egocentric biases were expected to be stronger when evaluating the facial expressions of similar targets than dissimilar targets.

The hypotheses, methods and statistical analyses of Study 4 were preregistered (https://osf.io/kjb84).

2. Study 1: Facial mimicry, reward & oxytocin

The influence of reward on facial mimicry: no evidence for a significant effect of oxytocin

Irene Trilla, Hanna Drimalla, Malek Bajbouj and Isabel Dziobek

Abstract. Recent findings suggest a role of oxytocin in the tendency to spontaneously mimic the emotional facial expressions of others. Oxytocin-related increases of facial mimicry, however, seem to be dependent on contextual factors. Given previous literature showing that people preferentially mimic emotional expressions of individuals associated with high (vs low) rewards, we examined whether the reward value of the mimicked agent is one factor influencing the oxytocin effects on facial mimicry. To test this hypothesis, 60 male adults received 24 IU of either intranasal oxytocin or placebo in a double-blind, between-subject experiment. Next, the value of neutral male faces was manipulated using an associative learning task with monetary rewards. After the reward associations were learned, participants watched videos of the same faces displaying happy and angry expressions. Facial reactions to the emotional expressions were measured with electromyography. We found that participants judged as more pleasant the face identities associated with high reward values than with low reward values. However, happy expressions by low rewarding faces were more spontaneously mimicked than high rewarding faces. Contrary to our expectations, we did not find a significant direct effect of intranasal oxytocin on facial mimicry, nor on the reward-driven modulation of mimicry. Our results support the notion that mimicry is a complex process that depends on contextual factors, but failed to provide conclusive evidence of a role of oxytocin on the modulation of facial mimicry.

This article has been published as:

Trilla, I., Drimalla, H., Bajbouj, M., & Dziobek, I. (2020). The Influence of Reward on Facial Mimicry: No Evidence for a Significant Effect of Oxytocin. *Frontiers in Behavioral Neuroscience*, 14. https://doi.org/10.3389/fnbeh.2020.00088

2.1 Introduction

Facial mimicry is defined as the automatic imitation of emotional facial expressions of others. It is an inherent aspect of social behaviour and acts as a social regulator by reinforcing social bonds and facilitating the understanding of others' emotional states (Hess & Fischer, 2013; Niedenthal, 2007). Facial mimicry is distinguished from other affective processes that may also lead to congruent facial reactions, such as emotional contagion or affective empathy, in that mimicked facial expressions reflect the sharing of the emotional displays, rather than a response to the other's emotional state (Hess & Blairy, 2001; Hess & Fischer, 2014).

Though initial theories understood mimicry as a stimulus-driven response, whereby the mere perception of facial expression would elicit a matching response in the observer, it is now well-established that mimicry depends on several factors related to the social context and the relationship between the interactants (Fischer & Hess, 2017). For example, pre-existing social bonds, goals to affiliate, similarity, positive mood, and a pro-social orientation have been shown to increase the tendency to mimic (see Seibt et al., 2015 for a review of social modulators of facial mimicry). These observations have motivated the notion of mimicry as a context-specific social process that occurs when there is a motivation to affiliate with the other person (Fischer & Hess, 2017) or when the interaction with the other would increase social wellbeing (Wang & Hamilton, 2012).

One of the factors shown to modulate mimicry is the reward value of the interactant. Using an implicit conditioning paradigm, Sims et al. (2012) associated different faces to losing and winning monetary rewards. Participants showed stronger facial mimicry in response to happy expressions displayed by faces previously conditioned with winning money compared to losing money, a finding that was replicated by Korb et al. (2019). In line with these results, fMRI and EEG studies using the same reward manipulation showed a greater functional coupling between reward- and mimicry-related brain areas in response to high rewarding faces (Sims et al., 2014), as well as stronger mu suppression, considered to be an index of cortical motor simulation (Trilla Gros et al., 2015). Further support for a link between reward and facial mimicry comes from a study in which sad faces were mimicked only when participants were monetarily rewarded for accurately identifying the emotional expression (Hess, Blaison, et al., 2017). In trials in which participants did not expect any reward, viewing sad faces elicited a smile instead, possibly indicative of a reaction of Schadenfreude. Altogether, these studies demonstrate that the reward value ascribed to the interactant influences the tendency to mimic their emotional expressions, arguably by impacting the implicit liking and motivation to affiliate with the person.

On the neurobiological level, oxytocin has been proposed as one endocrine factor, together with vasopressin and testosterone, influencing the modulation of mimicry (Kraaijenvanger et al., 2017). Oxytocin is a neuropeptide involved in several physiological and psychological

functions that regulate both social and non-social behaviour (Quintana & Guastella, 2020). Amongst other social processes, oxytocin has been shown to play a role in emotion recognition (Shahrestani et al., 2013) and empathy (Hurlemann et al., 2010). Given the role of mimicry in facilitating emotion understanding and regulating social behaviour, it seems reasonable to hypothesize that some socio-cognitive effects of oxytocin could be at least partly mediated by an influence on facial mimicry. In line with this, Korb et al. (2016) tested whether intranasal administrations of oxytocin would enhance facial mimicry while making emotion judgments. Oxytocin increased facial mimicry in response to infants' expressions of anger, but only marginally for adult targets. A small marginal increase was found for mimicry of infants' expressions of happiness. Although these results suggest some involvement of oxytocin in facial mimicry, the effects seem to depend on contextual factors such as the emotion and age of the interactant.

If the role of oxytocin is to promote adaptive social behaviour (Ma et al., 2016), we would expect that intranasal oxytocin enhances facial mimicry only in those contexts where there's a motivation to affiliate with the other. Based on this, Pavarini et al. (2019) investigated whether oxytocin increased mimicry of approachable emotions (e.g., happiness, sadness) more than non-approachable emotions (e.g., fear, anger). They found that intranasal administrations of oxytocin enhanced mimicry of sadness and happiness, although the latter only in individuals who showed low positive expressivity. Concerning non-approachable emotions, no significant effects of oxytocin were found for mimicry of anger. Though these results show tentative evidence that oxytocin may selectively increase mimicry for emotions that inspire social approach, the small effect size and the complexity of these results warrant further investigation into the contextual effects of oxytocin on facial mimicry.

The current study aimed to examine the influence of oxytocin on the reward modulation of facial mimicry. Based on the idea that oxytocin promotes adaptive social behaviour, and that facial mimicry preferentially occurs when we interact with rewarding others, we hypothesized that intranasal administrations of oxytocin would enhance facial mimicry compared to placebo, but more so when viewing emotional expressions of faces associated with higher reward value.

2.2 Methods

2.2.1 Participants

Sixty healthy men (M_{age} = 27.40, SD_{age} = 6.03) were recruited for this study. Only male participants were included to avoid gender differences in oxytocin response (e.g., Lynn et al., 2014; Rilling et al., 2014). All participants were German native speakers. Exclusion criteria included a history of psychiatric or neurological disorders, heart and cardiovascular conditions, other severe medical conditions (e.g., chronic pain syndrome, chronic degenerative or

inflammatory central nervous system diseases), substance use, current psychoactive medication, history of allergic or toxic reactions, smoking, participation in a pharmacological study in the last four months before the study, and nasal congestion or colds. Participants were asked to abstain from alcohol 24 h before the experiment, and from eating and drinking caffeine beverages for two h before the experiment.

All participants gave written informed consent and were financially remunerated for their participation. The study was conducted in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki, 6th revision), and was approved by the Ethics Committee of the Department of Psychology at Humboldt-Universität zu Berlin.

2.2.2 Procedure

The study followed a double-blind, placebo-controlled, between-subject design. At the start of the experimental session, participants filled out demographic information and completed the Positive and Negative Affect Schedule (PANAS; Krohne et al., 1996) to assess their current mood. Next, participants self-administered 24 IU of either oxytocin (n = 30) or placebo (n = 30) with a nasal spray. The intranasal administration was guided by the experimenter and followed the provider's indications (Apotheke des Universitätsklinikums Heidelberg, Germany), as well as recommendations by Guastella et al. (2013). Both the experimenter and the participant were blind to the content of the nasal spray, and the treatment assignment was done randomly.

Right after treatment administration, participants completed questionnaires of verbal intelligence (Mehrfachwahl-Wortschatz-Intelligenztest-B; MWT-B; Lehrl, 2005), empathy (Empathy Quotient; EQ; Simon Baron-Cohen & Wheelwright, 2004) and autistic traits (Autism Spectrum Quotient, 33-items German version; AQ; Simon Baron-Cohen et al., 2001; Freitag et al., 2007). Because the physiological effects of intranasal oxytocin do not begin until 30 min post-administration (Spengler et al., 2017) and questionnaire completion took a maximum of 15 min, we assumed that participants' answers would not be affected by the treatment. During the waiting time until the next mood assessment, participants watched an affectively neutral documentary unrelated to the content of the experiment about the prehistoric monument of Stonehenge.

Thirty-five minutes after treatment administration, participants' mood was assessed again with the PANAS. Next, participants completed a reward learning paradigm in which different neutral faces were associated with low and high reward values. After the face-reward associations were learned, participants watched videos of the same faces displaying happy and angry expressions (facial mimicry task). Electromyography (EMG) was used to track the participants' facial expressions while watching emotion displays. Facial EMG is a widely used method for measuring facial mimicry as it allows to detect subtle face reactions that may be undetectable visually (van Boxtel, 2010). The main experimental tasks began between

39 and 50 min (M = 42.6, SD = 2.2) after treatment administration, and continued for roughly 60 min.

At the end of the session, participants completed a short questionnaire assessing blinding integrity and were debriefed about the study aims. Participants were notified about the content of the nasal spray after the data collection of the full sample was finalized.

The reward learning task was programmed in MATLAB R2015b (The MathWorks, Inc., Natick, MA, USA) using the Psychophysics Toolbox extension (Kleiner et al., 2007). OpenSesame (Mathôt et al., 2012) was used for stimulus presentation for the mimicry task. Questionnaires were implemented online using the software package SoSci Survey (Leiner, 2018).

2.2.3 Reward learning task

An associative learning paradigm based on Valentin and O'Doherty (2009) was used to pair neutral faces with low and high reward values. Pictures of four male faces with neutral expressions were selected from a validated emotion expressions database (Kliemann et al., 2013) and presented in pairs. In each trial, participants had to choose between one of the two faces displayed side by side (picture size: 400 × 400 px) by pressing either the right or left arrow key on a keyboard. Upon selecting a face, participants could either win 10 cents of a Euro (win trial) or nothing (no-win trial). For each of the two stimulus pairs, one face was associated with a 60% probability of winning (high reward condition), and the other face with a 30% probability of winning (low reward condition). The assignment of face identities to reward conditions was randomized across participants. Participants were not disclosed about the exact reward probabilities assigned to each face but were informed that one face would lead to a higher number of win trials overall, compared to the other face in the pair. After the participant's response, the outcome ("10 cents" or "0 cents"), as well as the accumulated earnings, were shown for 2000 ms (see Figure 2.1). Participants were instructed to maximize their earnings. A fixation cross was displayed at the beginning of each trial for a variable duration between 500 and 1500 ms.

The task consisted of a minimum of three blocks of 20 trials (10 trials per face pair, presented in random order). Participants completed up to four additional 10-trial blocks for each face pair if the proportion of high reward choices did not reach 80% by the end of each block. This learning criterion was set to make sure that the assimilation of the face-reward associations was comparable across all participants and stimuli. If participants did not reach the 80% criterium by the end of the task, the corresponding face stimuli were considered not learned and were excluded from the EMG analyses (see "EMG Analysis" section). On average, participants completed 4.20 blocks (SD = 1.44) per face pair. There were no significant differences in the number of blocks completed between oxytocin and placebo groups, $t(58) = 0.14$, $p = 0.89$.

To check whether the reward associations changed how faces were evaluated, participants were asked to rate the pictures for pleasantness using a 7-point scale (1: very unpleasant; 7: very pleasant) before and after the reward learning task.

2.2.4 Facial mimicry task

A passive viewing task was used to assess facial mimicry. In each trial, participants watched a 4000 ms-video of a face displaying either a happy or an angry expression (768 × 768 px). Videos started with a neutral face that changed into the emotional expression, which peaked at around 1500 ms. Dynamic expressions were used as they have been shown to elicit stronger facial mimicry responses than static pictures and are more ecologically valid (Sato et al., 2008). Preceding each video, a fixation cross was presented for 1500 ms.

In total, participants watched happy and angry facial expressions of six male actors. Four of the face identities had been previously associated with either high or low reward values in the reward learning task. The remaining two identities were new to the participants (unconditioned faces) and were used as a control condition to assess the direct influence of oxytocin on facial mimicry. All videos were presented eight times (96 trials total) in randomized order.

To assure that participants paid attention to the videos, in 25% of the trials a yes/no question about the actors' physical attributes (e.g., presence of a beard, hair colour) was asked right after the clip presentation (see Figure 2.1). During the mimicry task, videos of the participants were recorded with a webcam to detect potential artefacts in the EMG data (e.g., if participants sneezed, moved, et cetera).

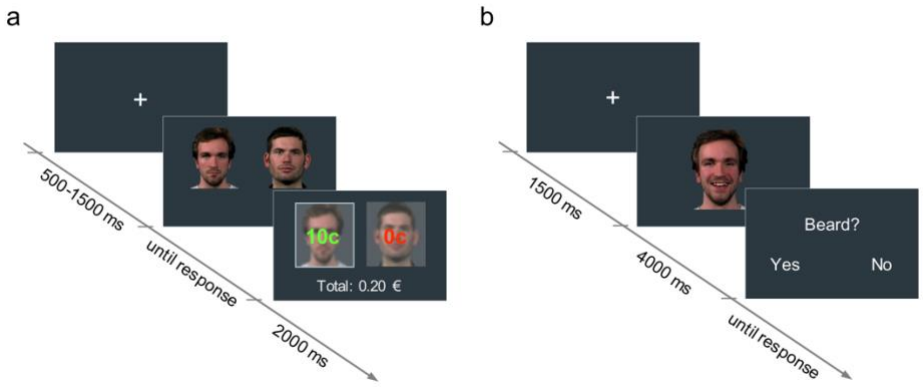

Figure 2.1. (a) Reward learning task: in each trial, participants had to choose between two neutral faces. One face was associated with a 60% probability of winning 10 cents (high reward condition), and the other face with a 30% probability of winning 10 cents (low reward condition). The trial outcome (i "10 cents" in green, or "0 cents" in red, for the win and no-win trials, respectively) was displayed superimposed on each picture for 2000 ms. The accumulated earnings were shown below. (b) Facial mimicry task: in each trial, participants watched a 4000-ms video of a face displaying a happy or an angry expression. In 25% of the trials, participants had to answer an attention-control question concerning a physical attribute of the face they had just seen.

2.2.5 EMG data acquisition

During the facial mimicry task, EMG was used to record the activity of the zygomaticus major (ZM), a muscle on the corners of the mouth that is activated when smiling, and the corrugator supercilii (CS), a muscle located in the eyebrow area that contracts when frowning. Stronger activity in the ZM compared to CS is commonly used as an index of mimicry of happiness, while the opposite pattern (i.e. higher CS than ZM) reflects an expression congruent with anger (Hess, Arslan, et al., 2017; van Boxtel, 2010). Bipolar Ag/AgCl electrodes were attached to the left side of the face over the two muscles. The ground electrode was placed on the centre of the forehead, below the hairline. We followed standard EMG site preparation and electrode placement procedures (van Boxtel, 2010). Skin conductance electrode gel was used to facilitate conductance between the electrodes and the facial skin. To cover the recording of muscular activity, participants were told that facial electrodes were measuring sweat production.

EMG signals were amplified with EMG amplifiers (Becker Meditec, Karlsruhe, Germany; gain = 1230; frequency response 19–500 Hz). The amplified signals were digitized using a USB multifunction card USB-6002 (National Instruments Inc., Ireland) connected to a laptop computer Dell Latitude 5540, running data acquisition software DasyLab 10.0 (National Instruments Ireland Resources Limited). The raw EMG signals were sampled with 500 Hz and 16-bit resolution. Within DasyLab, signals were online (RMS root mean square) integrated with a time constant of 50 ms and rectified. The multifunction card USB-6002 acquired also trigger signals from the parallel port of the presentation computer. For further processing, the integrated EMG and the trigger signals were down-sampled to 20 Hz and stored as an ASCII file.

2.2.6 EMG data reduction and artefact control

EMG data was pre-processed offline in Matlab R2015b (The MathWorks, Inc., Natick, MA, USA) using self-made scripts. Data were segmented from 1500 ms before to 4000 ms after stimulus onset. The raw EMG signal and the video recordings of all participants were screened visually to detect trials with artefacts. Artefacts were defined as distortions in the EMG data associated with, for example, resting the chin on the hand, swinging the head, yawning, eye closing, mumbling, and displaying facial expressions in the pre-stimulus period. On average, 3.79 trials (SD = 3.56) were rejected per participant. There were no significant differences in the number of trials rejected by emotion and reward conditions, $F(2, 114) = 0.35$, $p = 0.71$. Artefact-free data were Z-standardized within muscles and within participants to account for individual and muscle differences. Due to technical issues during EMG acquisition, data in the 500 ms directly before and after stimulus onset were distorted and could not be used. The period from 1500 ms to 4000 ms after stimulus onset was used as the window of interest. As in Sims et al. (2012); this interval was determined based on the

time when the emotion expressions peaked during the videos (i.e. at around 1500 ms in this study), and because it included the period of maximal EMG responses (see Figure 2.3 and Figure 2.4). Change from baseline scores were calculated for each trial and muscle by subtracting the mean EMG amplitude from 500 to 1000 ms preceding stimulus onset (baseline) from the mean EMG amplitude of the window of interest. The resulting baseline-corrected EMG scores were used as the dependent variable in the statistical analyses.

2.2.7 Statistical analyses

Data and code necessary to reproduce the analyses reported here are available at https://osf.io/n85sh. All statistical analyses were run in R (R Core Team, 2020) and R studio (RStudio Team, 2019). The main R packages used were: *afex* for ANOVA (Singmann et al., 2019); *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017) and *emmeans* (Lenth, 2020) for linear mixed-effects analysis; *pwr* (Champely, 2018) and *TOSTER* (Lakens, 2017) for equivalence testing; and *ggplot2* (Wickham, 2009) for figures.

2.2.7.1 Reward manipulation check

To test whether participants learned the reward associations, the proportion of high reward choices for each of the first three learning blocks were fitted in a 2 × 2 mixed ANOVA with Block (first, second, third) as within-subject factor and Treatment (placebo, oxytocin) as between-subject factor. A 2 × 2 × 2 mixed ANOVA with Reward (low reward, high reward) and Time (pre-, post-reward learning task) as within-subject factors, and Treatment (placebo, oxytocin) as between-subject factor, was used to examine whether face pleasantness ratings changed after the reward associations were learned. The Holm-Bonferroni method was applied to adjust for multiple comparisons in post hoc tests.

2.2.7.2 EMG analysis

Due to technical problems, EMG data of two participants were not collected. Data from two additional participants were excluded from the EMG analyses due to low accuracy (less than 80%) on the attention control questions in the mimicry task. The EMG analysis sample thus included data from 29 participants in the oxytocin group and 27 participants in the placebo group.

EMG data were analysed using linear mixed models (LMM). Separate LMMs were executed to test EMG responses to observing happy and angry expressions. Mimicry of happiness was defined as increased activation of ZM vs CS response to happy faces. The opposite pattern of muscular activity (i.e. stronger CS vs ZM response) in response to angry faces would indicate mimicry of anger.

First, we examined the direct influence of oxytocin on facial mimicry. EMG data from trials in which participants viewed unconditioned faces (i.e. faces that did not appear in the reward learning task, and thus had not been associated with any particular reward value) were used as the dependent variable in LMMs with Muscle (ZM, CS), Treatment (placebo, oxytocin) and their interaction as fixed effects.

Second, to test the effects of reward on facial mimicry, and the modulatory role of oxytocin, we fitted into LMMs the EMG data from trials that presented low reward and high reward faces. Trials in which the face presented did not reach the learning criterion by the end of the reward learning task were excluded. The LMMs included the main effects of Muscle (ZM, CS), Reward (low reward, high reward), and Treatment (placebo, oxytocin), as well as the corresponding two- and three-way interactions, as fixed effects.

To account for non-independencies in the data, we entered by-participant and by-stimulus random intercepts in all LMMs. If a model led to singular fits, the random-intercept for stimuli was removed. Sum to zero contrasts were set for all predictors. *P*-values were computed based on Satterthwaite approximation for denominator degrees of freedom.

2.3 Results

Overall, 37% of the participants correctly guessed the treatment they had received, 40% made an incorrect guess, and 23% reported not knowing. The proportion of participants who made correct or incorrect guesses did not significantly differ between the placebo and oxytocin groups, $\chi 2(2) = 1.49$, $p = 0.47$.

With respect to mood changes, participants in both the placebo and oxytocin groups reported less negative affect after treatment administration ($M = 12.6$, $SD = 3.97$) than at baseline ($M = 14.02$, $SD = 4.62$), as shown by a significant main effect of time, $F(1, 58) = 14.35$, $MSE = 4.19$, $p < 0.001$, $\eta^2_G = 0.027$. No significant change in positive affect was observed, nor a significant effect of treatment for neither positive nor negative affect (all $p > 0.47$). Descriptive statistics of questionnaire scores are available in the supplementary Table S2.1.

2.3.1 Learning of reward associations

In line with the expected reward learning curve, the proportion of high reward choices significantly increased with the number of blocks, $F(1.99, 115.52) = 36.51$, $MSE = 0.01$, $p < 0.001$, $\eta^2_G = 0.10$ (Figure 2.2). Participants chose the high reward faces more often in the second ($M = 0.72$, $SD = 0.21$), $t(116) = -6.33$, $p < 0.001$, Cohen's $d = -0.60$, and third blocks ($M = 0.75$, $SD = 0.20$), $t(116) = -8.14$, $p < 0.001$, Cohen's $d = -0.80$, compared to the first block of trials ($M = 0.60$, $SD = 0.19$). The difference between the second and third block did not reach statistical significance, $t(116) = -1.80$, $p = 0.07$, Cohen's $d = -0.17$. There was no significant main effect of treatment, $F(1, 58) = 0.26$, $MSE = 0.10$, $p = 0.61$, $\eta^2_G = 0.004$, nor a significant interaction

with block, $F(1.99, 115.52) = 0.16$, $MSE = 0.01$, $p = 0.85$, $\eta^2_G < 0.001$, which is consistent with previous studies in which intranasal oxytocin did not enhance learning with non-social reinforcements compared to placebo (Clark-Elford et al., 2014; Hurlemann et al., 2010). Descriptive statistics on the proportion of high reward choices for each block and treatment group can be found in the Supplementary Table S2.2.

A change in the perceived pleasantness of the faces after the reward learning task, indicated by a significant reward-by-time interaction, $F(1, 58) = 8.74$, $MSE = 0.31$, $p = 0.004$, $\eta^2_G = 0.01$, supports that face-reward associations were assimilated (Figure 2.2). Planned simple effects confirmed that, while there were no significant differences in pleasantness ratings at baseline (low reward: $M = 3.72$, $SD = 0.71$; high reward: $M = 3.77$, $SD = 0.89$), $t(94.30) = -0.36$, $p > 0.99$, Cohen's $d = -0.06$, faces associated with high reward value were rated as more pleasant ($M = 4.23$, $SD = 0.98$) than faces associated with low reward value ($M = 3.76$, $SD = 0.91$) after the reward learning task, $t(94.30) = -3.37$, $p = 0.003$, Cohen's $d = -0.50$. Treatment group did not significantly interact with these effects, $F(1, 58) = 0.41$, $MSE = 0.31$, $p = 0.53$, $\eta^2_G = 0.001$, nor had an overall influence on pleasantness ratings, $F(1, 58) = 1.16$, $MSE = 1.52$, $p = 0.29$, $\eta^2_G = 0.010$. See Supplementary Table S2.3 for descriptive statistics on the pleasantness ratings for each treatment group.

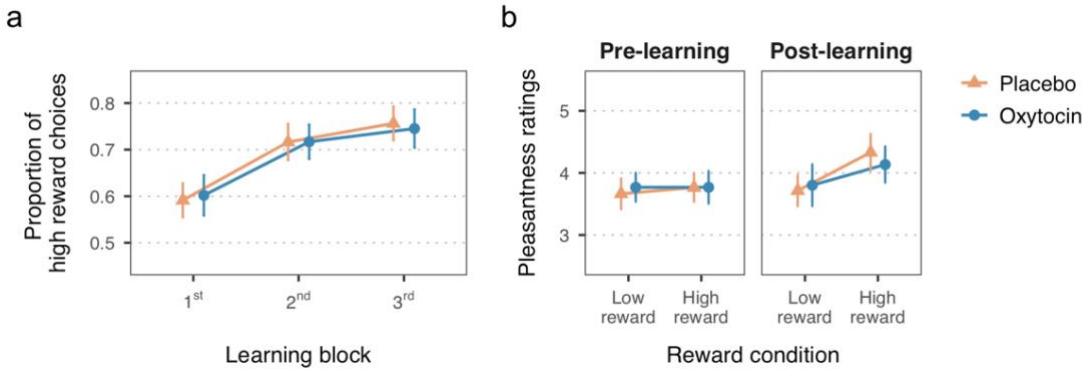

Figure 2.2. (a) Mean proportion of high reward choices for the first three learning blocks of the reward learning task. (b) Mean pleasantness ratings for the faces paired with low and high reward values, before and after the reward learning task. Error bars are within-subject 95% confidence intervals.

2.3.2 Effects of oxytocin on facial mimicry

Observation of happy facial expressions elicited higher ZM ($M = 0.05$, $SE = 0.04$) compared to CS activity ($M = -0.06$, $SE = 0.04$), $b = 0.05$, $SE = 0.02$, 95% $CI = (0.02, 0.09)$, $t = 3.04$, $p = 0.002$, confirming the occurrence of mimicry of happiness (Figure 2.3). Contrary to our expectations, the interaction between muscle and treatment was not significant, $b = -0.03$, $SE = 0.02$, 95% $CI = (-0.06, 0.01)$, $t = -1.65$, $p = 0.10$. The main effect of treatment, however, was

significant, b = 0.06, SE = 0.02, 95% CI = (0.02, 0.10), t = 2.71, p = 0.007, indicating that both ZM and CS were more activated in the oxytocin group (M = −0.06, SE = 0.04) compared to the placebo group (M = 0.05, SE = 0.04).

With respect to responses to angry facial expressions, we did not find any significant difference between ZM and CS, b = 0.01, SE = 0.02, 95% CI = (−0.03, 0.06), t = 0.64, p = 0.52, which indicates that our task did not elicit mimicry of anger (Figure 2.3). No significant main effect of treatment, b = −0.003, SE = 0.03, 95% CI = (−0.05, 0.05), t = −0.11, p = 0.91, nor an interaction with muscle reactivity were found, b = 0.02, SE = 0.02, 95% CI = (−0.03, 0.06), t = 0.82, p = 0.41.

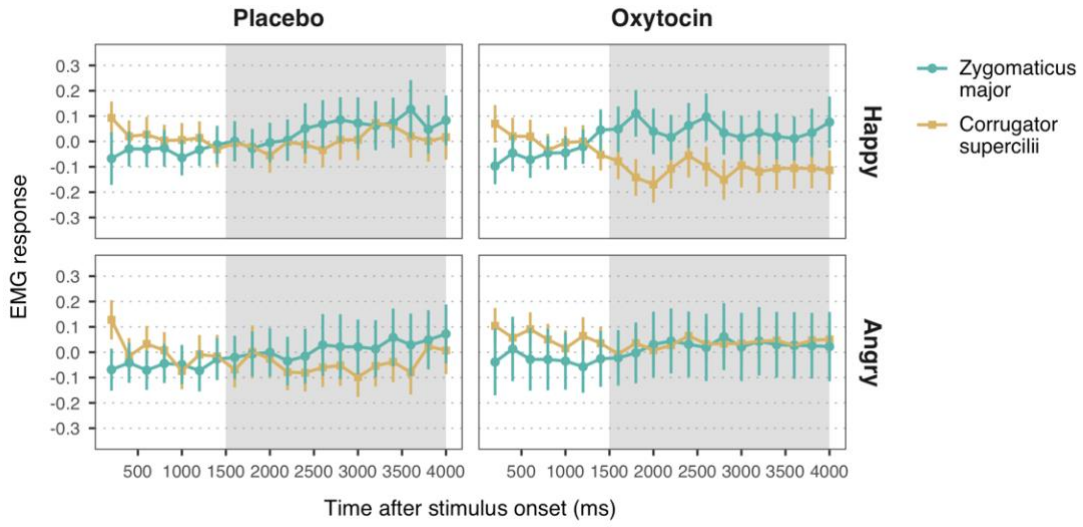

Figure 2.3. Zygomaticus major and corrugator supercilii responses to happy and angry facial expressions of unconditioned faces, for each treatment group. Plotted data corresponds to Z-standardized, baseline-corrected electromyographic (EMG) activity averaged within 200-ms time-bins. Only data from the time-window between 1500 and 4000 ms post-stimulus onset (shaded in grey) was used for statistical analysis. Error bars represent within-subject 95% confidence intervals.

2.3.3 Effects of reward on facial mimicry and the influence of oxytocin

The LMM with EMG responses to high and low reward happy faces as dependent variable yielded a significant main effect of muscle, b = 0.06, SE = 0.01, 95% CI = (0.03, 0.08), t = 4.37, p < 0.001. As before, viewing happy facial expressions elicited stronger ZM activity (M = 0.07, SE = 0.03) than CS activity (M = −0.05, SE = 0.03), confirming mimicry of happiness. Moreover, we found a significant muscle-by-reward interaction, b = 0.03, SE = 0.01, 95% CI = (0.002, 0.05), t = 2.10, p = 0.036. Follow-up analyses on the estimated means showed that ZM response was higher for low reward happy faces (M = 0.12, SE = 0.03) compared to high reward faces (M = 0.02, SE = 0.03), t(2552.70) = 2.54, p = 0.03 (Figure 2.4). This result indicates that, contrary to our expectations, low reward happy faces elicited stronger mimicry than

higher reward happy faces. No significant differences in CS activity were found between high and low reward faces, $t(2552.70) = -0.42$, $p = 0.68$. Neither the predicted three-way interaction, nor any of the two-way interactions with treatment reached significance (all $p > 0.39$). These results thus provide no significant evidence for a modulation of oxytocin on the effects of reward on mimicry of happiness.

The LMM on EMG responses to angry expressions did not yield any significant main effects or interactions (all $p > 0.14$; Figure 2.4). See the supplementary tables for the descriptive statistics (Table S2.4) and the complete results of the LMMs (Table S2.5, Table S2.6).

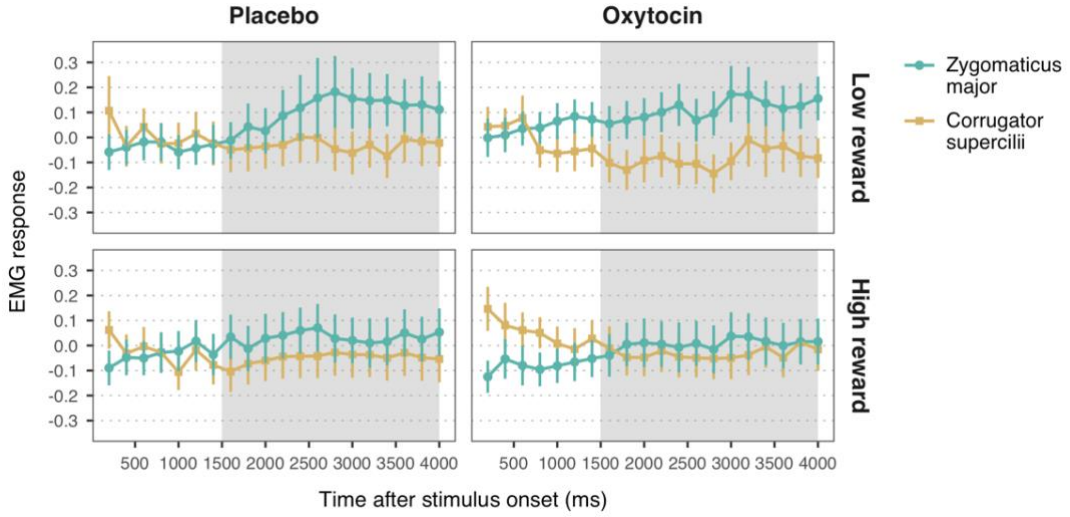

Figure 2.4. Zygomaticus major and corrugator supercilii responses to happy faces associated with low and high reward values, for each treatment group. Plotted data corresponds to Z-standardized, baseline-corrected EMG activity averaged within 200-ms time-bins. Only data from the time-window between 1500 and 4000 ms post-stimulus onset (shaded in grey) was used for statistical analysis. Error bars represent within-subject 95% confidence intervals.

2.3.4 Equivalence testing

Because null-hypothesis significance testing can only reject the presence of an effect, we cannot conclude that intranasal oxytocin does not influence facial mimicry based on the non-significant results reported above. Using the Two One-Sided Tests (TOST) procedure of equivalence testing (Lakens, 2017), we re-examined our main null findings to test whether the effect of oxytocin on facial mimicry was statistically equivalent to the placebo effect, or whether our data was just not sensitive enough to detect the predicted group differences. We limited this analysis to EMG responses to happy faces, as our task failed to elicit mimicry of anger. Given that there is no clear theoretical boundary for oxytocin's socio-cognitive effects for setting the equivalence bounds (Quintana, 2018), we defined the smallest effect size of interest-based on the smallest effect size detectable with 80% power given our sample size (Quintana, 2018; Tabak et al., 2019). The alpha level was set to 0.05.

First, we applied the TOST procedure to test the null effect of oxytocin on EMG responses to happy expressions of unconditioned faces. Because the TOST procedure is based on t-tests, we reduced the original Muscle-by-Treatment interaction tested in the LMM by computing mimicry indices (i.e. difference score between CS and ZM activity in response to happy faces) for each participant. Positive scores indicate the occurrence of mimicry of happiness. The equivalence test comparing the mean mimicry index of the oxytocin vs placebo groups was non-significant, $t(49.15) = 1.55$, $p = 0.06$, given equivalence bounds of −0.76 and 0.76. This indicates that we cannot reject effects bigger than what could be reliably tested based on the statistical power of our study.

Second, we tested the oxytocin effects on the influence of reward on mimicry of happiness. To reduce the original three-way interaction in the LMM (Muscle × Reward × Treatment), we subtracted the mimicry index to high reward faces from the mimicry index to low reward faces for each participant. Positive scores indicate stronger mimicry of happiness in response to high reward vs low reward faces. The equivalence test comparing the oxytocin and placebo groups did not reach significance, $t(48.98) = -1.64$, $p = 0.05$, given equivalence bounds of −0.80 and −0.80. Note that the equivalence bounds were recalculated for this second TOST because the sample size for this analysis was smaller: data from two participants of the oxytocin group and 3 of the placebo group were not included because they did not reach the learning criterium for any of the face pairs in the reward learning task.

Based on the equivalence tests and the null-hypothesis tests combined, we can neither conclude that oxytocin has an effect on facial mimicry nor reliably reject effect sizes that could be detected with 80% power given the sample size of this study.

2.4 Discussion

This study sought to investigate the modulatory effect of intranasal oxytocin on the link between reward and facial mimicry. We hypothesized that intranasal administration of oxytocin would increase facial mimicry, but more so in response to faces previously associated with high reward compared to low reward value.

Our study failed to confirm our original hypotheses. First, we found an influence of reward on facial mimicry of happiness, but this effect was in the opposite direction as predicted: happy faces associated with low reward were mimicked more than happy faces associated with high reward. Second, we did not find evidence for a significant effect of oxytocin on facial mimicry, neither for a direct influence, nor on the modulation of the reward-mimicry link.

2.4.1 The influence of reward on facial mimicry

Previous studies reported a higher tendency to mimic the emotional expressions of faces paired with high rewards than with low rewards (Hess, Blaison, et al., 2017; Korb et al., 2019; Sims et al., 2012). Here, we manipulated the value of different neutral faces using an associative learning task with monetary rewards. No differences in the learning of face-reward associations were found between the oxytocin and placebo groups, in line with previous studies in which intranasal oxytocin did also not facilitate learning with non-social reinforcements (Clark-Elford et al., 2014; Hurlemann et al., 2010). As expected, our reward manipulation changed how participants evaluated the faces, such that those linked with higher monetary rewards were rated as more pleasant. However, the EMG results did not replicate the direction of the reward effects of previous mimicry studies, as participants showed stronger mimicry in response to happy faces previously conditioned with low reward compared to high rewarding happy faces.

Differences in the sample characteristics and experimental design may account for this unexpected result. For example, previous studies on the reward-mimicry link had predominantly female samples, and their stimuli included both female and male targets (Hess, Blaison, et al., 2017; Korb et al., 2019; Sims et al., 2012). Conversely, we only included male participants and same-gender face identities. Given that there are gender differences in how characteristics of the target person, such as gaze direction and gender, influence mimicry reactions (Hess & Bourgeois, 2010; Schrammel et al., 2009), it could be that the direction of reward effects is partly dependent on the gender of and the relationship between the expresser and the perceiver.

Our study also differed from previous ones in how reward was manipulated. In Hess, Blaison, et al. (2017), the motivation to affiliate, rather than the value of the targets themselves, was manipulated by either rewarding (or not rewarding) the correct identification of their emotional expressions. In contrast, Sims et al. (2012) and the current study aimed to directly alter the value of the faces, although the paradigm and reward schemas used were different. First, Sims et al. (2012) used a classical conditioning task to pair different reward values with face identities implicitly. While this task proved to be effective in manipulating the value of the stimuli (Korb et al., 2019), it had the limitation that the participant's learning of the associations could not be tracked, as task performance was unrelated to the associations between the face stimuli and monetary rewards. In contrast, we implemented a reward learning task in which the reward outcomes were contingent on the participant's choices. This allowed us to use task performance as an indication of the learning of the face-reward pairings.

Second, Sims et al. (2012) compared mimicry responses to faces associated with rewards (winning money) vs faces associated with punishments (losing money). In the current study, face stimuli were only conditioned with different probabilities of winning money. We chose relatively low reward probabilities (60% for high reward condition vs 30% for low reward

condition) to keep the learning implicit, as pilot testing suggested that at these rates, participants were not aware that they had learned the associations, despite learning curves and pleasantness ratings were as expected. Note that, even though reward probabilities were overall lower than those used in Sims et al. (2012), faces assigned to the high reward condition were anyway paired twice as often with a monetary outcome than faces in the low reward condition.

These variations in the reward schemas may have affected how the manipulation influenced the evaluation of the faces compared to previous studies. Rather than associating a positive or negative valence to the face identities (e.g., targets paired with positive outcomes vs targets paired with negative outcomes), our manipulation may have changed the level of uncertainty as to which each face would be accompanied with a rewarding outcome. According to a recent account of social cognition, the inability to precisely predict the states and actions of others during social interactions is associated with aversive feelings (FeldmanHall & Shenhav, 2019). On this account, processes of inference and affect-sharing such as emotional mimicry are thought to be activated to reduce social uncertainty. Therefore, participants may have reacted with stronger mimicry to low reward faces to promote emotion understanding, regain a bond with the interactant, and ultimately reduce the higher social uncertainty that the low reward faces conveyed compared to faces associated with higher reward probability.

2.4.2 Null effects of intranasal oxytocin

Contrary to our predictions, intranasal administrations of oxytocin did not influence the degree of facial mimicry of happiness, nor the modulation by reward. Two studies have previously investigated the effects of oxytocin on facial mimicry, and neither found intranasal oxytocin to consistently change the facial reactivity to happy expressions (Korb et al., 2016; Pavarini et al., 2019). Only in their exploratory analysis did Pavarini et al. (2019) detect oxytocin-related increases in mimicry reactions to happy faces in a subsample of participants who showed reduced positivity expressivity at baseline. A meta-analysis found small increases in the expression of positive emotions after intranasal oxytocin, although it did not significantly improve the recognition of happiness (Leppanen et al., 2017). Due to the sample size, our study had limited power to detect small or medium effects, a shortcoming that was further highlighted by the inconclusive results of the equivalence tests applied. While our data could not provide conclusive evidence on the (lack of) effects of intranasal oxytocin, cumulative null findings may indicate that, even if oxytocin would influence the degree of mimicry of happiness, this effect is probably small.

Mimicry of happiness is highly frequent and very consistent across individuals and social contexts (Bourgeois & Hess, 2008; Hess & Bourgeois, 2010), which may render an effect of oxytocin difficult to detect. Instead, oxytocin manipulations may be more prone to affect

mimicry reactions to emotions that are more context- and person-dependent, such as sadness and anger. Given that our mimicry task did not elicit congruent facial responses to angry expressions, we were unable to test the effects of oxytocin and reward on mimicry for anger. However, previous studies did find oxytocin-related increases in mimicry in response to angry expressions (Korb et al., 2016) and a trend towards an increase of mimicry of sadness (Pavarini et al., 2019).

Nevertheless, current evidence of the role of oxytocin on facial mimicry remains weak. Beyond our null findings, oxytocin did not have a robust direct influence on mimicry reactions in previous studies. Rather, effects were only observed in response to certain stimuli (e.g., children's emotional faces vs adult faces, in Korb et al., 2016), emotional expressions (e.g., angry vs happy in Korb et al., 2016; and sad vs happy and angry, in Pavarini et al., 2019), and type of mimicry tasks (e.g., offset task vs intensity task, in Korb et al., 2016). Altogether, this suggests that oxytocin may not act by directly influencing the basic mechanisms underlying mimicry. Instead, it could play a role in the modulation of facial mimicry by the social context. This would fit well with current theories of oxytocin, which posit that oxytocin improves social adaptation by increasing social salience (Shamay-Tsoory & Abu-Akel, 2016) and promoting approach behaviour and reducing avoidance (Harari-Dahan & Bernstein, 2014). However, our results on the oxytocin modulation of the reward-mimicry link failed to provide conclusive evidence supporting this hypothesis. Further research is therefore warranted to better disentangle under which conditions intranasal oxytocin has a reliable effect.

2.4.3 Limitations

One limitation of our study is that we were not able to test oxytocin effects on mimicry of anger. Given that an affiliation intent is necessary for emotional mimicry (Fischer & Hess, 2017), and that angry expressions may be interpreted as a threat signal, our task may not have provided a suitable context for mimicry of anger to occur. To test whether oxytocin effects on mimicry are emotion-specific, future studies should present additional emotional expressions, and measure facial reactions in more interactive paradigms that promote affiliative goals.

Even though the sample size of this study was similar to other oxytocin studies (e.g., Korb et al., 2016), it had insufficient power to detect small-to-medium effects, which are the effect sizes typically observed in oxytocin research (Walum et al., 2016). As evidenced by the inconclusive results from the equivalence tests, further studies with higher statistical power are needed to draw reliable conclusions about the presence or absence of a meaningful effect of oxytocin on facial mimicry. Prospective research would benefit from determining a-priori the smallest effect size of interest for equivalence testing. Even though setting the equivalence boundaries according to the sample size of this study helped us to determine that our

data was not sensitive enough to detect even large effects, a better approach would derive more precise and meaningful theoretical predictions based on prior studies, instead of on a resource question (Lakens et al., 2018). Also, the inclusion of female participants in future studies is encouraged, as so far, all intranasal oxytocin studies on mimicry have exclusively tested men, and gender differences are commonly observed in oxytocin research (e.g., Lynn et al., 2014; Rilling et al., 2014).

Finally, while intranasal oxytocin is the most accessible and widespread method to study the role of oxytocin on human behaviour, this methodology is not without limitations. First, even though oxytocin is administered intranasally to take advantage of the direct pathways between the nasal cavity and the central nervous system (Quintana et al., 2015), it is still unclear what doses are needed to reach relevant brain regions and induce behavioural effects (Leng & Ludwig, 2016). Compared to 12 IU and 40 IU, the dose administered here (24 IU) exerted the maximum impact on neural reactivity in a study on dose-dependency (F. B. Spengler et al., 2017). However, dose-dependent effects were not observed on the behavioural level. Moreover, it is yet to be determined to what extend oxytocin's central actions account for its socio-cognitive effects, or whether the peripheral oxytocinergic system is also involved (Leng & Ludwig, 2016; Quintana et al., 2015; Valstad et al., 2016). Further research is thus warranted to improve our understanding of the oxytocinergic system and to develop improved methods to study the role of oxytocin on social behaviour and cognition.

2.4.4 Conclusion

Results from this study add to the evidence that facial mimicry is influenced by the reward value of the interactant and reinforces the notion of mimicry as a context-specific social process. Nevertheless, the fact that the reward effects were in the opposite direction as reported in previous mimicry studies highlights the need to closely evaluate the impact of experimental protocols, sample characteristics, and contextual factors on the modulation of mimicry by reward. Other replication attempts in psychological science show that small differences in study designs may lead to meaningful changes in the results (Noah et al., 2018). Rather than invalidating the reported findings, failed replications should be taken as an opportunity to identify new possible moderators of the investigated effects (Van Bavel et al., 2016).

Also, we did not find evidence for a significant role of oxytocin on the effect of reward on mimicry in response to happy expressions. While this is the first time, to our knowledge, that the influence of oxytocin on the reward-mimicry link is investigated, our results coincide with previous reports of null effects of oxytocin on mimicry of happiness. Nevertheless, the sample size of this study, which limited the power to detect small or medium effects, and the fact that only male participants were included, warrant a cautious interpretation and generalizability of these null results. As in other fields, oxytocin research has suffered from

strong publication bias (Lane et al., 2016), and some of the early findings have not been replicated (e.g., Nave et al., 2015). Parallel to efforts in improving the methodological quality of oxytocin studies (Walum et al., 2016), it is warranted that failed replications and null results like this one are brought to light and considered when assessing the role of oxytocin on social cognitive processes.

2.5 Supplementary tables

Table S2.1. Mean (standard deviation) of questionnaire scores for the full sample and for each treatment group.

Measurements		Full sample (n = 60)	Placebo (n = 30)	Oxytocin (n = 30)
Autism Spectrum Quotient		8.40 (4.69)	8.87 (4.01)	7.93 (5.31)
Empathy Quotient		36.33 (10.49)	36.10 (11.08)	36.57 (10.04)
Mehrfachwahl-Wortschatz-Intelligenztest-B		26.67 (4.58)	27.33 (4.82)	26.00 (4.32)
Positive affect scale (PANAS)	Pre-treatment	30.98 (6.91)	30.27 (7.70)	31.70 (6.08)
	Post-treatment	30.47 (6.95)	29.97 (7.49)	30.97 (6.46)
Negative affect scale (PANAS)	Pre-treatment	14.02 (4.62)	14.13 (5.08)	13.90 (4.20)
	Post-treatment	12.60 (3.97)	12.73 (4.77)	12.47 (3.05)

Note. PANAS: Positive and Negative Affect Schedule.

Table S2.2. Mean (standard deviation) proportion of high reward choices in each of the first three trial blocks of the reward learning task.

	Full sample (n = 60)	Placebo (n = 30)	Oxytocin (n = 30)
Block 1	0.60 (0.19)	0.59 (0.19)	0.60 (0.19)
Block 2	0.72 (0.21)	0.70 (0.21)	0.73 (0.22)
Block 3	0.75 (0.20)	0.73 (0.21)	0.77 (0.19)

Table S2.3. Mean pleasantness ratings (standard deviation) for the faces paired with low and high reward values, as rated before and after the reward learning task.

		Full sample (n = 60)	Placebo (n = 30)	Oxytocin (n = 30)
Pre-reward learning	Low reward	3.72 (0.71)	3.75 (0.76)	3.68 (0.66)
	High reward	3.77 (0.89)	3.85 (0.78)	3.68 (0.99)
Post-reward learning	Low reward	3.76 (0.91)	3.80 (0.95)	3.72 (0.88)
	High reward	4.23 (0.98)	4.42 (0.81)	4.05 (1.12)

Table S2.4. Mean (standard deviation) EMG scores for each recorded muscle, as a function of treatment group, reward and emotion conditions. EMG scores refer to Z-standardized, baseline-corrected EMG data for the interval between 1500 and 4000 ms post-stimulus onset.

		Placebo (n = 27)		Oxytocin (n = 29)	
		Zygomaticus major	Corrugator supercilii	Zygomaticus major	Corrugator supercilii
Happy expressions	No reward	0.02 (1.08)	-0.03 (0.86)	0.08 (1.10)	-0.08 (0.93)
	Low reward	0.11 (0.97)	-0.04 (0.76)	0.12 (0.73)	-0.08 (0.71)
	High reward	0.03 (0.85)	-0.06 (0.74)	0.02 (0.85)	-0.03 (0.70)
Angry expressions	No reward	0.04 (1.38)	-0.02 (0.88)	0.005 (1.75)	0.01 (0.78)
	Low reward	0.04 (0.85)	0.0001 (0.62)	0.05 (0.75)	-0.02 (0.81)
	High reward	-0.04 (0.76)	-0.01 (0.87)	0.03 (0.96)	-0.06 (0.76)

Table S2.5. Results of the linear mixed models conducted to test the effect of treatment (placebo vs oxytocin) on EMG scores for each recorded muscle (zygomaticus major vs corrugator supercilii) in response to expressions of happiness (Model 1) and anger (Model 2) by unconditioned faces (i.e. faces not presented in the reward learning task).

	Model 1: EMG responses to happy expressions						Model 2: EMG responses to angry expressions					
	Estimate	*SE*	*95% CI*	*ß*	*t*	*p*	*Estimate*	*SE*	*95% CI*	*ß*	*t*	*p*
Muscle	0.05	0.02	0.02 – 0.09	0.07	3.04	.002	0.01	0.02	-0.03 – 0.06	0.02	0.64	.522
Treatment	0.06	0.02	0.02 – 0.10	0.08	2.71	.007	-0.003	0.03	-0.05 – 0.05	-0.003	-0.11	.910
Muscle x Treatment	-0.03	0.02	-0.06 – 0.01	-0.04	-1.65	.100	0.02	0.02	-0.03 – 0.06	0.02	0.82	.414
Random Effects												
σ^2	0.51						0.82					
τ_{00}	0.01 participant, 0.003 stimuli						0.01 participant					
ICC	0.02						0.01					
N	56 participant, 6 stimuli						56 participant					
Observations	1734						1706					
Marginal R^2/ Conditional R^2	0.013 / 0.035						0.001 / 0.013					

Note. *p*-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Significant *p*-values, according to alpha < 0.05, are indicated in bold. Model equation: `EMG scores ~ Muscle + Treatment + Muscle*Treatment + (1|Participant) + (1|Stimuli)`. EMG scores refer to the Z-standardized, baseline-corrected EMG activity for each muscle, averaged within the interval between 1500 and 4000 ms post-stimulus onset. The inclusion of the random intercept for stimuli in Model 2 led to singular fits and was therefore removed from the final model. Estimate: unstandardized coefficient; SE: standard error; CI: confidence interval; *ß*: standardized coefficient; *t*: test statistic coefficient; *p*: p-value; σ^2: within-group variance; τ_{00} = between-group variance; ICC = interclass correlation (ratio of between-cluster variance to total variance); N: number of random effects.

Table S2.6. Results of the linear mixed models conducted to test the effect of treatment (placebo vs oxytocin) on EMG scores for each recorded muscle (zygomaticus major vs corrugator supercilii) in response to expressions of happiness (Model 3) and anger (Model 4) by faces previously associated with low reward and high reward.

	Model 3: EMG responses to happy expressions						Model 4: EMG responses to angry expressions					
	Estimate	*SE*	*95% CI*	*ß*	*t*	*p*	*Estimate*	*SE*	*95% CI*	*ß*	*t*	*p*
Muscle	0.06	0.01	0.03 – 0.08	0.08	4.37	<.001	0.02	0.01	-0.01 – 0.05	0.03	1.48	.140
Reward	0.02	0.01	-0.01 – 0.05	0.03	1.50	.134	0.02	0.01	-0.01 – 0.04	0.02	1.26	.209
Treatment	0.03	0.02	-0.02 – 0.07	0.04	1.24	.215	-0.004	0.02	-0.03 – 0.03	-0.01	-0.28	.777
Muscle x Reward	0.03	0.01	0.00 – 0.05	0.04	2.10	.036	0.01	0.01	-0.02 – 0.03	0.01	0.53	.594
Muscle x Treatment	-0.002	0.01	-0.03 – 0.02	-0.002	-0.12	.902	-0.02	0.01	-0.05 – 0.01	-0.03	-1.43	.152
Reward x Treatment	0.005	0.01	-0.02 – 0.03	0.01	0.36	.719	0.004	0.01	-0.02 – 0.03	0.01	0.28	.777
Muscle x Reward x Treatment	-0.01	0.01	-0.04 – 0.01	-0.02	-0.85	.397	0.01	0.01	-0.02 – 0.04	0.02	0.85	.397
Random Effects												
σ^2	0.53						0.53					
τ_{00}	0.01 $_{\text{participant}}$, 0.0004 $_{\text{stimuli}}$						0.003 $_{\text{participant}}$, 0.0003 $_{\text{stimuli}}$					
ICC	0.03						0.01					
N	51 $_{\text{participant}}$, 6 $_{\text{stimuli}}$						51 $_{\text{participant}}$, 6 $_{\text{stimuli}}$					
Observations	2966						2944					
Marginal R^2/ Conditional R^2	0.010 / 0.035						0.002 / 0.008					

Note. *p*-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Significant *p*-values according to alpha < 0.05 are indicated in bold. Model equation: `EMG scores ~ Muscle + Reward + Treatment + Muscle*Reward + Muscle*Treatment + Reward*Treatment + Muscle*Reward*Treatment + (1|Participant) + (1|Stimuli)`. EMG scores refers to the Z-standardized, baseline-corrected EMG activity for each muscle, averaged within the interval between 1500 and 4000 ms post-stimulus onset. Estimate: unstandardized coefficient; SE: standard error; CI: confidence interval; *ß*: standardized coefficient; *t*: test statistic coefficient; *p*: p-value; σ^2: within-group variance; τ_{00} = between-group variance; ICC = interclass correlation (ratio of between-cluster variance to total variance); N: number of random effects.

3. Study 2: Automatic imitation, gaze & autistic traits

Conditional effects of gaze on automatic imitation: the role of autistic traits

Irene Trilla, Hannah Wnendt and Isabel Dziobek

Abstract. Establishing direct gaze has been shown to enhance the tendency to automatically imitate the other person's actions, an effect that seems to be reduced in autism. Most previous studies, however, used experimental tasks that may have confounded the measurement of automatic imitation with spatial compatibility effects. This calls into question whether gaze cues regulate automatic imitation or instead affect domain-general processes of response inhibition. Using a task that disentangled imitative from spatial compatibility effects, the current study re-examined the role of autistic traits on the modulation of automatic imitation by direct and averted gaze cues. While our results do not provide evidence for an overall significant influence of gaze on neither automatic imitation nor spatial compatibility, autistic traits were predictive of reduced inhibition of imitative behaviour following averted gaze. Nonetheless, exploratory analyses suggested that the observed modulation by autistic traits may actually be better explained by the effects of concomitant social anxiety symptoms. In addition, the ethnicity of the imitated agent was identified as another potential modulator of the gaze effects on automatic imitation. Overall, our findings highlight the contextual nature of automatic imitation but call for a reconsideration of the role of gaze on imitative behaviour.

This article has been published as:

Trilla, I., Wnendt, H., & Dziobek, I. (2020). Conditional effects of gaze on automatic imitation: The role of autistic traits. *Scientific Reports, 10*(1), 15512.
https://doi.org/10.1038/s41598-020-72513-6

3.1 Introduction

Imitating the body postures, gestures and facial expressions of others is known to facilitate the understanding of their mental states (Drimalla et al., 2019) and to regulate social behaviour (Hess & Fischer, 2013; Lakin & Chartrand, 2003; Wang & Hamilton, 2012). While imitation may be intentional under certain circumstances, people tend to unconsciously copy observed body movements, even when they are irrelevant and could interfere with the task at hand, a process referred to as automatic imitation (Heyes, 2011). Automatic imitation has been shown to be an adaptive and flexible behaviour that is highly dependent on the social context. According to the social top-down response modulation (STORM) theory, automatic imitation is subtly controlled by social goals to promote one's social advantage (Wang & Hamilton, 2012).

One of the initial studies on the social modulation of automatic imitation identified gaze as an important signal that regulates imitative behaviour (Wang, Newport, et al., 2011). Together with facial expressions and body gestures, gaze cues are a rich source of non-verbal information to decode others' intentions and internal states during social interactions (Hamilton, 2016). Establishing direct gaze may signal social interest and intention to engage with the perceiver (Wirth et al., 2010), but it can also generate the feeling of being observed, an effect thought to promote prosocial behaviour (Hamilton & Lind, 2016). Moreover, gaze is important for social referencing, as it shifts the attentional focus toward the gaze direction and elicits joint attention (Frischen et al., 2007).

In the context of automatic imitation, Wang et al. (2011) showed that people tend to copy observed irrelevant movements more if the imitated agent establishes direct gaze, compared to when the agent averts the gaze away from the participant. This finding has since been replicated in successive experiments (Wang & Hamilton, 2014), which further demonstrated that the enhancement of automatic imitation following direct gaze is related to audience effects and the signalling of affiliation intent, rather than due to gaze-triggered shifts in spatial attention. Behavioural findings are supported by neuroscientific studies showing that direct gaze enhances neural mirroring of others' motor actions compared to averted gaze (Prinsen et al., 2017; Prinsen & Alaerts, 2019), and which identified the medial prefrontal cortex as a key brain region mediating the control of automatic imitation by gaze (Wang, Ramsey, et al., 2011).

While at first glance the evidence of gaze-triggered control of imitation seems strong, all aforementioned studies used an experimental paradigm that may have confounded the measurement of automatic imitation with more general processes of response inhibition, namely spatial compatibility. Typically, automatic imitation is assessed with stimulus-response compatibility tasks in which participants are required to perform a hand or finger movement (e.g., 'lift index finger') while at the same time observing a compatible (e.g., 'index finger

lift') or incompatible (e.g., 'middle finger lift') action by another agent (Heyes, 2011). A tendency to automatically imitate the observed irrelevant movement is indicated if the participant's performance is facilitated by observing a compatible action and/or interfered when observing an incompatible action. In many versions of this paradigm, however, the hand stimulus is displayed as a mirror view of the participant's hand, such that the observed movement is spatially aligned with the action required by the participant. In such conditions, the effects attributed to automatic imitation could also be explained in terms of the spatial compatibility between both movements (Catmur & Heyes, 2011): the participant's response might be facilitated not (only) because they both perform topographically similar movements (i.e. imitative compatibility), but because the stimulus and response actions involve a similar change in relative position (i.e. spatial compatibility). While many automatic imitation studies may have confounded both effects, there is evidence that imitative and spatial compatibility are two dissociable and independent processes (Boyer et al., 2012; Catmur & Heyes, 2011).

In an attempt to avoid spatial effects, the direction of the observed movement in the original automatic imitation paradigm with gaze cues was orthogonal to the response movement (Wang, Newport, et al., 2011). However, results from a recent meta-analysis indicate that orthogonal set-ups are not free from the influence of spatial compatibility effects (Cracco, Bardi, et al., 2018). Thus, this finding casts doubts on whether the gaze effects measured with this paradigm are actually affecting processes of automatic imitation or, instead, are influencing spatial compatibility.

With this confounder in mind, Marsh et al. (2016) examined the influence of gaze and group membership on automatic imitation using a paradigm that disentangled the effects of imitative and spatial compatibility. As in typical automatic imitation tasks, participants were asked to perform finger movements when observing imitatively compatible or incompatible actions. However, in half of the trials, participants observed finger movements by the agent's left hand (i.e. mirror view), and the other half presented the agent's right hand. While imitative and spatial compatibility coincide in the left-hand trials, the inclusion of right-hand trials allows dissociating automatic imitation from spatial compatibility. Using this set-up, Marsh et al. (2016) failed to replicate the effect of gaze on automatic imitation. Instead, social cues selectively influenced spatial compatibility, such that stronger spatial effects were found for in-group members with direct gaze and out-group members with averted gaze. Their results thus challenge conclusions from previous research on the social modulation of automatic imitation that did not control for the independent contributions of imitative and spatial processes.

In the face of these conflicting findings, the first aim of this preregistered study was to re-examine the influence of gaze cues on automatic imitation. As in the study by Marsh et al. (2016), we used a paradigm that measured the effect of direct and averted gaze cues on

automatic imitation and spatial compatibility independently. Based on the previous literature, we expected that imitative compatibility effects would be stronger following direct gaze than averted gaze. Alternatively, gaze could selectively affect spatial compatibility, such as found in Marsh et al. (2016). In this case, we would expect that direct gaze increases spatial compatibility effects compared to averted gaze.

Second, we tested whether autistic traits modulate how gaze cues influence automatic imitation. A previous study found that individuals with autism spectrum conditions (ASC) tend to automatically imitate others' actions, but the strength of their imitative responses is not regulated according to the gaze direction of the observed agent (Forbes et al., 2017). Importantly, this reduced contextual modulation of imitation could not be completely attributed to insensitivity to gaze cues, as direct gaze elicited alerting responses in both ASC and control samples. This and other evidence of atypical imitative behaviour in response to social signals (Cook & Bird, 2012; Forbes et al., 2017) may seem in conflict with studies showing intact automatic imitation in individuals with autistic traits (Cracco, Bardi, et al., 2018; Darda et al., 2020). In an attempt to integrate these findings, motivational accounts such as the STORM theory have proposed that, although the basic mechanisms of imitation seem to be preserved, individuals with ASC may be impaired in adjusting their imitative behaviour to the social context (Wang & Hamilton, 2012). Accordingly, our second hypothesis predicted that the impact of gaze cues on automatic imitation would be weaker with increasing levels of autistic traits in a non-clinical sample.

In addition to the prespecified research questions, we explored if the influence of gaze on automatic imitation is conditional to other contextual factors. This question was motivated by earlier research showing that gaze cues are processed differently depending on the ethnicity and group membership of the observed face (Collova et al., 2017; Marsh et al., 2016; Weisbuch et al., 2017). For example, in the study by Marsh et al. (2016), in-group members with direct gaze elicited stronger compatibility effects than out-group members with direct gaze. Given these context-dependent reactions to gaze, we examined whether the ethnicity of the observed agent would also shape the impact of gaze on automatic imitation.

Lastly, we explored the role of social anxiety as a predictor of the impact of gaze on automatic imitation. Social anxiety is highly prevalent in individuals with ASC, and is also characterized by impairments in social attention, such as fear and avoidance of direct eye contact (Hessels et al., 2018; Kleberg et al., 2017). Despite the similarities in some of the social impairments, atypical gaze patterns in social anxiety seem to be more consistent with anxiety-driven avoidance, while alterations in gaze responding in ASC have been related to reduced social motivation (Kleberg et al., 2017; Kliemann et al., 2010). Examining whether and how social anxiety symptoms modulate the influence of gaze on automatic imitation could shed light on the mechanisms behind the observed effects in ASC.

To sum up, the current study aimed to (1) re-test the influence of direct and averted gaze on automatic imitation using a task that disentangled imitative and spatial compatibility, and (2) examine whether autistic traits modulate the influence of gaze on compatibility effects. Exploratory analyses further examined whether other social cues (i.e. ethnicity of the observed agent) and individual differences in social functioning (i.e. social anxiety) shape the impact of gaze on imitative behaviour.

3.2 Methods

The preregistration form of this study is available at https://osf.io/84wqe.

3.2.1 Participants

Sixty participants (31 females, 28 males, 1 non-binary; M_{age} = 26.6, SD_{age} = 4.80; 58 right-handed, 2 ambidextrous) took part in this study. An a-priori power analysis using G*Power 3 (Faul et al., 2007) estimated a sample of 35 to 63 participants (alpha = 0.05, power = 0.9, within-subject repeated-measures analysis of variance) for a η_p^2 reported between .35 (Gaze x Automatic imitation effect in Experiment 1; Wang, Newport, et al., 2011)) and .23 (Gaze x Automatic imitation effect in the control sample; Forbes et al., 2017).

None of the participants reported current psychiatric or neurological disorders, current psychoactive medication, or history of regular substance use. Only participants with normal or corrected-to-normal vision and who made fewer than three errors in the Ishihara test for colour-blindness (Ishihara, 1972) were included.

All participants gave written informed consent and were financially remunerated for their participation. The study was conducted in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and was approved by the Ethics Committee of the Psychology department at Humboldt-Universität zu Berlin.

3.2.2 Materials

3.2.2.1 Stimuli

Clips of five male and five female actors performing direct and averted head movements were selected from the Amsterdam Dynamic Facial Expression Set (ADFES; van der Schalk et al., 2011). In the direct gaze clips, the actors start with the head oriented to their left side and turn towards the observer. In the averted gaze clips, the actors start facing the observer and turn the head towards their right. The original videos were cut to 2500 ms, and copies with vertically flipped frames were additionally created to obtain direct and averted gaze clips in which the head moves towards the left. The size of the clips was 768 x 576 pixels. Actors maintained a neutral expression in all clips. Two of the identities (1 male, 1 female)

were presented in the practice trials, and the remaining eight identities were used for the experimental trials. For each gender, half of the actors were of Northern-European descent (white), and half were Mediterranean (Turkish or Moroccan; dark-skinned).

Gaze clips were combined with hand stimuli (Figure 3.1). Different female and male right hands were photographed to create the frames that would be sequentially presented in the automatic imitation task to simulate index and middle finger movements. In total, seven frames were obtained for each hand: the hand in resting position on a vertical panel, the hand with the index finger fully lifted, the hand with the middle finger fully lifted, and two intermediate positions for each of the two finger movements. Hand images were paired with the gaze stimuli based on the physical attributes of the actors and hands. All hand images were edited to match the size and skin tone to each actor and were flipped to obtain both right- and left-hand stimuli.

3.2.2.2 Automatic imitation task

A stimulus-response compatibility paradigm was used to measure the effects of gaze on imitative and spatial compatibility. Participants were required to lift the index or middle finger of their right hand in response to colour cues, while at the same time observing imitatively congruent (e.g., index finger lift when prompted to lift the index finger) or incongruent actions (e.g., middle finger lift when prompted to lift the index finger). The irrelevant finger movements were performed by actors who had previously directed the gaze towards the participant (direct gaze condition) or away from them (averted gaze condition). To dissociate imitative and spatial compatibility effects, half of the trials presented the actors' left hand, and the other half showed their right hand (Catmur & Heyes, 2011). With this set-up, imitative and spatial compatibility overlapped in the left-hand trials (i.e. mirror view), but these effects were disentangled in the right-hand trials (Figure 3.1).

Each trial of the task started with a fixation cross located at the height of the actors' eye area for 1000 ms (Figure 3.2). Next, participants observed a 2500-ms clip of a direct or an averted head movement, with the actor's hand in resting position. To avoid anticipatory responses, the final frame of the gaze clip with the resting hand remained static for a variable duration selected randomly between 200 and 800 ms. Three finger movement frames were then presented sequentially for 34 ms to induce the apparent motion of an index or middle finger lift. Eighty milliseconds after the movement onset, a purple or a green dot appeared superimposed between the actor's index and middle finger knuckles to cue the participant's required response. The stimulus-response correspondence (green dot = "lift index finger", purple dot = "lift middle finger", or vice versa) was counterbalanced across participants. An asynchronous onset of the response cue with respect to the irrelevant finger movement has been shown to facilitate imitative compatibility effects (Catmur & Heyes, 2011). The last frame of the finger movement remained onscreen until a response was made, or after 2000 ms.

3.2.2.3 Autistic traits and social anxiety questionnaires

The Autism Spectrum Quotient (AQ) was used to measure individual differences in autistic traits (Baron-Cohen et al., 2001). The short German version of the AQ (Freitag et al., 2007) consists of 33 items assessing different aspects related to ASC (e.g., social and communication skills, imagination and attention). The AQ has been shown to have good test-retest reliability and inter-rater reliability (Baron-Cohen et al., 2001) as well as good discriminative validity (Woodbury-Smith et al., 2005).

Social anxiety was assessed with the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) and the Social Phobia Scale (SPS; Mattick & Clarke, 1998). These are widely used self-report scales that evaluate two categories of feared situations in social anxiety disorders: those related to being observed by others (SPS), and those related to social interaction (SIAS). Finally, the German version of the Gaze Anxiety Rating Scale (GARS) was used to measure self-reported fear and avoidance of making eye contact in social situations (Domes et al., 2016).

The internal consistency of all questionnaires in our sample was high, as indexed by Cronbach's $\alpha > .82$. Descriptive statistics for all questionnaire scores, Cronbach's α and the correlations between autistic traits and social anxiety measures are available in the supplementary Table S3.8.

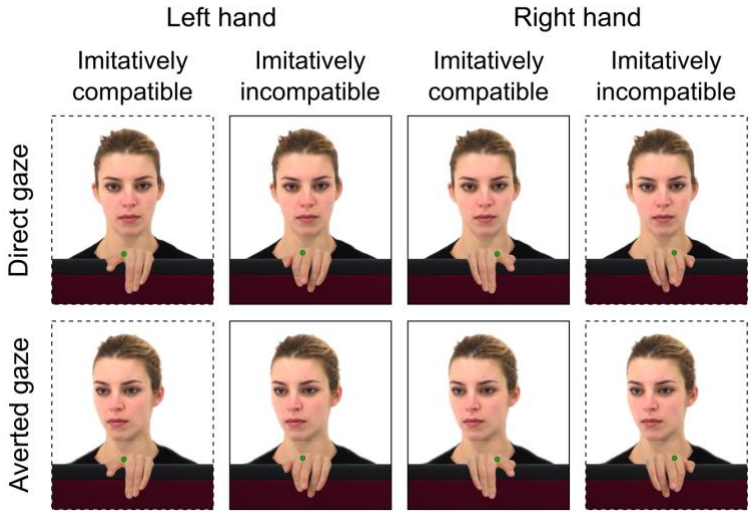

Figure 3.1. Experimental conditions of the automatic imitation task. Three factors were manipulated following a full factorial within-subject design: Gaze (direct, averted), Imitative compatibility (compatible, incompatible) and Spatial compatibility (compatible, incompatible). Dashed frames indicate spatially compatible conditions; solid frames indicate spatially incompatible conditions. In this example, participants would be required to lift their right index finger in response to a green dot, and the right middle finger in response to a purple dot.

3.2.3 Procedure

The testing session started with the automatic imitation task. MATLAB R2016b (The MathWorks, Inc., Natick, Massachusetts, United States) and the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007) were used to present the stimuli and collect the responses. Participants were instructed to pay attention to the head movement of the actors, and to perform the corresponding finger action at the appearance of the response cue. They were encouraged to respond as fast as possible without sacrificing accuracy. Throughout the task, participants held down two keyboard keys ("n" and "m", marked with a green and a purple sticker) with the index and middle finger of their right hand. The key released upon making a finger lift indicated the participant's response. Reaction times were recorded from the onset of the response cue until the key release. To ensure spatial (in)compatibility between the observed and performed movements, the position of the participant's hand was matched to the orientation of the hand presented on the screen by placing the keyboard vertically at a 45° angle (Figure 3.2).

After the instructions, participants completed 10 practice trials with accuracy feedback to train the mapping between the colour of the response cue and the finger responses. If participants made more than two errors, additional trials were performed until a cumulative accuracy of 70% was reached. On average, participants completed 10.08 (*SD* = 0.33) practice trials.

The experimental phase consisted of 256 trials (32 trials per condition) without feedback, divided into 4 blocks. Each block comprised 64 trials, balanced for the colour of the response cue (green, purple), imitative compatibility (compatible, incompatible), the observed finger movement (index finger lift, middle finger lift) and the identity of the actor. Across all blocks, trials were also balanced for the hand presented (right hand, left hand) and the direction of the head movement of the gaze clips (towards the right, towards the left). All trials within a block presented the same hand of the actors. The order of hand blocks ('*right-left-right-left*' or '*left-right-left-right*') was counterbalanced across participants. Trials within blocks were randomly ordered. Participants could take a short break between blocks.

After the automatic imitation task, participants were asked to rate on a 5-point scale (0 = not at all, 4 = a lot) whether they generally felt "observed", "ignored", "connected", "confirmed", "rejected", "put under pressure", "relieved" and "distracted" when watching the clips of the actors directing the gaze towards them, and the clips of the actors directing the gaze away from them. These ratings aimed to assess the meaning that participants attributed to direct and averted gaze signals. Lastly, participants provided basic demographic information (age, gender, occupation and level of education) and completed the questionnaires assessing autistic traits and social anxiety, which were implemented in SoSci Survey (Leiner, 2018).

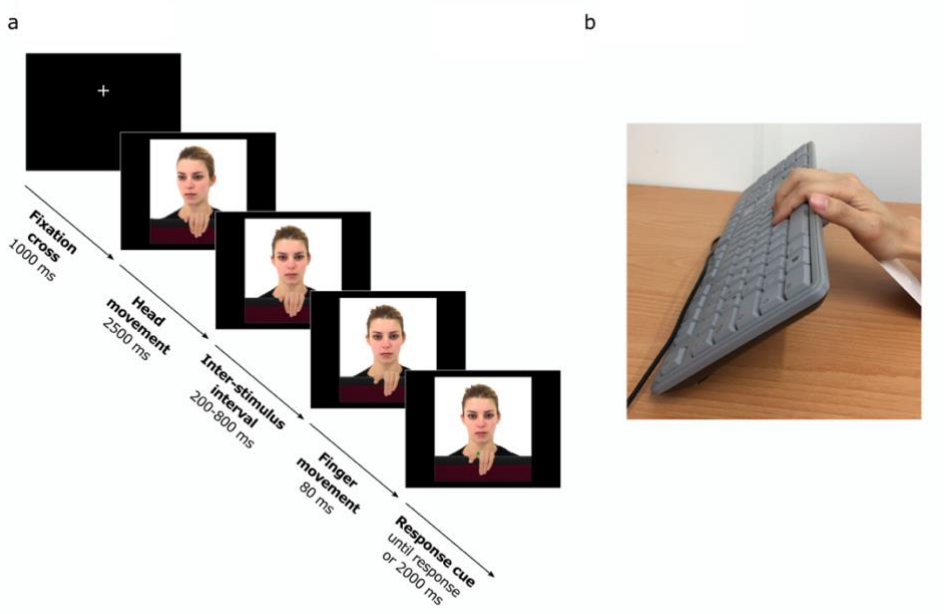

Figure 3.2. (a) Trial sequence of the automatic imitation task. This example represents a direct gaze, left-hand trial. (b) Keyboard set-up used to match the position of the participant's hand to the orientation of the hand stimuli displayed.

3.2.4 Statistical analysis

All statistical analyses were conducted in R (R Core Team, 2020) and R studio (RStudio Team, 2019). Data and code necessary to reproduce all analyses reported here are available at https://osf.io/9gku6/.

3.2.4.1 Data exclusion

Reaction times (RT) and error rates were assessed independently as measures of performance. For RT analyses, trials were excluded if no response was made, if the response was incorrect, or if RTs were smaller than 200 ms or greater than 1000 ms. To minimize the effect of outliers, trials that deviated 1.5 times from the interquartile range of RTs within each condition and participant were further excluded. We chose the interquartile range instead of the standard deviation to measure dispersion for outlier detection as it is more robust against extreme values and non-normality. On average, 8.55% (*SD* = 7.41) of the trials were excluded per participant. For error rate analyses, only trials in which no response was made within the 2000 ms-response window were excluded. This led to a rejection of 0.05% (*SD* = 0.18) of trials per participant.

3.2.4.2 Confirmatory analyses

Generalized linear mixed-effects models (GLMM) on single-trial data were used to test our hypotheses. Our original plan was to conduct analyses of variance on aggregate data (i.e. on

the mean RT and error rates for each condition and participant) to follow the statistical methods used in previous studies (see preregistration at https://osf.io/84wqe). However, GLMM provides additional advantages as they allow to: (1) account for random effects for both participants and stimuli, so more of the error can be modelled, (2) fit the data of individual trials instead of the means for each participant, which gives more statistical power, (3) accommodate missing data, so we could include data of all participants irrespective of the number excluded trials, and (4) test the effect of continuous variables, such as AQ scores (Kliegl et al., 2011). Moreover, GLMMs allowed us to specify the distribution of the dependent variable to match the distributional properties of raw RT and error rate data (Lo & Andrews, 2015). This facilitates interpretability and comparison of results to previous studies that analysed untransformed RT. Unless otherwise specified below, results from the GLMMs reported here led to the same conclusions as the preregistered analyses, which can be found in the supplementary material.

The GLMMs on RT and error rates data included the following fixed effects: main effects of 'Imitative compatibility' (2 levels: compatible, incompatible), 'Spatial compatibility' (2 levels: compatible, incompatible), 'Gaze' (2 levels: direct, averted) and 'AQ' (continuous variable); the 2-way interactions 'Gaze*Imitative', 'Gaze*Spatial', 'AQ*Imitative', 'AQ*Spatial', and 'AQ*Gaze'; and the 3-way interactions 'AQ*Gaze*Imitative' and 'AQ*Gaze*Spatial'.

To simplify the pattern of results, we restricted the description of results to: (1) the main effects of imitative and spatial compatibility, which would indicate the occurrence of compatibility effects, and the main effect of gaze, which would suggest that gaze cues were processed; (2) the two-way interaction 'Gaze*Imitative' and 'Gaze*Spatial, which would show that the degree of the corresponding compatibility effect depends on the gaze direction of the observed agent; and (3) the 3-way interactions 'AQ*Gaze*Imitative' and 'AQ*Gaze*Spatial', which would suggest that autistic traits modulate the influence of gaze on imitative (or spatial) compatibility.

Fixed effects with categorical predictors were tested using effect coding contrasts, and continuous predictors (i.e. AQ) were mean-centred. To account for non-independencies in the data, all GLMMs included by-participant and by-stimulus random intercepts. The levels of the factor 'Stimulus' corresponded to each of the possible combinations of the actor presented in the gaze clips, the direction of their head movement (turn towards the right side, turn towards the left side), the finger that the actor lifted (index finger, middle finger) and the colour of the response cue (green, purple).

GLMMs that included RT as the dependent variable used an Inverse Gaussian distribution with Identity link function. The Inverse Gaussian is a right-skewed unimodal distribution with continuous responses greater than or equal to 0 that reproduces the distributional shape of raw RT (Lo & Andrews, 2015). This was chosen over other possible distributions (i.e. Gaussian and Gamma) on the basis of model comparisons using AIC and BIC values and

likelihood-ratio tests. The Identity link function was selected as we assumed that our manipulations linearly affected the RT rather than some function of RT (Lo & Andrews, 2015). This assumption underlies experiments based on mental chronometry and is inherent when using linear regression or linear mixed models. A mixed-effect logistic regression with binomial distribution was conducted for error rate data to adhere to the binary nature of this variable (correct response, incorrect response).

For all models, *p*-values were calculated using Wald-statistics approximation. The statistical threshold was set at $p < .05$, and tests were two-tailed.

3.2.4.3 Exploratory analyses

Exploratory analyses were conducted on RT data, as this was a more sensitive measure than error rates. The statistical procedure applied for each exploratory analysis is described in the corresponding Results subsection to ease comprehensibility.

3.3 Results

3.3.1 Confirmatory analyses

The mean and standard deviation of RT and error rates for each condition, as well as the complete GLMM statistics, are presented in the supplementary Table S3.2 and Table S3.3.

3.3.1.1 Reaction times

The GLMM on RT data confirmed the occurrence of both imitative compatibility, $b = 10.23$, 95% *CI* = [7.57, 12.90], $SE = 1.36$, $t = 7.53$, $p < .001$, and spatial compatibility effects, $b = 22.31$, *95% CI* = [19.63, 24.99], $SE = 1.37$, $t = 16.34$, $p < .001$ (Figure 3.3). That is, participants were faster to perform correct finger movements when they observed an imitatively compatible ($M = 489.44$, $SE = 1.40$) than an imitatively incompatible action ($M = 501.46$, $SE = 1.54$). Similarly, participants responded faster to a spatially compatible action ($M = 485.56$, $SE = 1.52$) compared to a spatially incompatible action ($M = 505.53$, $SE = 1.40$). The main effect of gaze was also statistically significant, $b = -4.62$, *95% CI* = [-7.29, -1.96], $SE = 1.36$, $t = -3.41$, $p = .001$, indicating that participants responded faster following direct gaze ($M = 492.95$, $SE = 1.48$) than after averted gaze ($M = 497.83$, $SE = 1.47$).

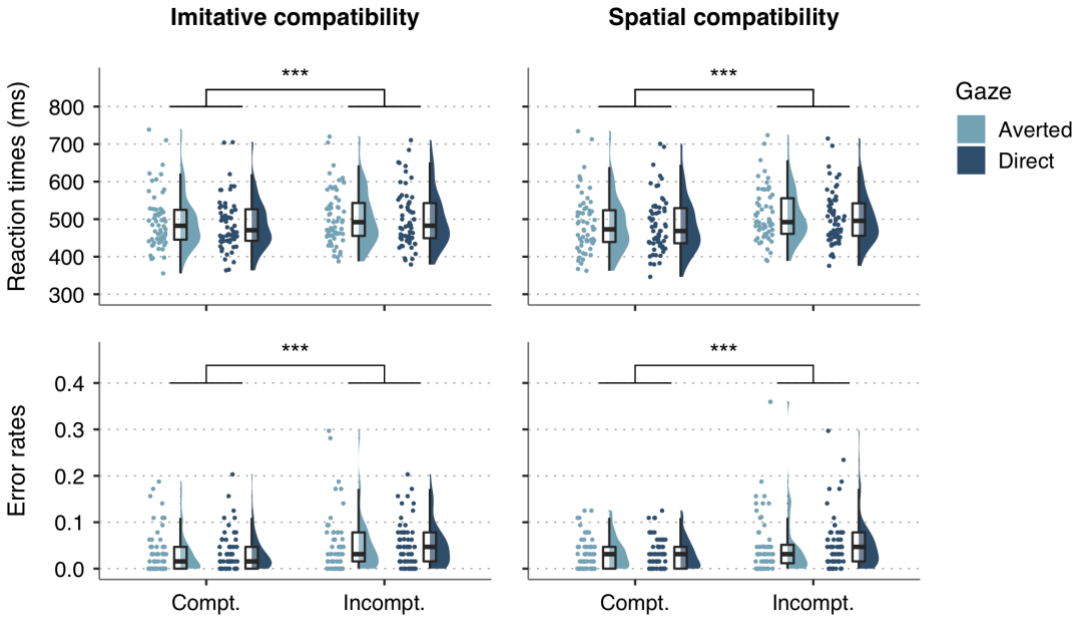

Figure 3.3. Raincloud plots of reaction times (upper panels) and error rates (lower panels) of imitatively compatible and incompatible trials (left panels), and of spatially compatible and incompatible trials (right panels), for both direct gaze (dark blue) and averted gaze (light blue) conditions. Each point represents the average reaction time or error rate of an individual participant for the corresponding condition. *** $p < .001$.

Gaze did not significantly interact with neither imitative compatibility, $b = 2.51$, *95% CI* = [-2.75, 7.78], $SE = 2.69$, $t = 0.93$, $p = .35$, nor spatial compatibility, $b = -5.10$, *95% CI* = [-10.44, 0.23], $SE = 2.72$, $t = -1.87$, $p = .06$. The 3-way interaction between spatial compatibility, gaze and AQ was also not significant, $b = 0.16$, *95% CI* = [-0.83, 1.14], $SE = 0.50$, $t = 0.31$, $p = .75$. However, AQ significantly modulated the effect of gaze on imitative compatibility, $b = -1.29$, *95% CI* = [-2.28, -0.31], $SE = 0.50$, $t = 2.59$, $p = .01$. To further disentangle this 3-way interaction, we tested the conditional effects of AQ on imitative compatibility for direct gaze and averted gaze conditions separately. As seen in Figure 3.4, the imitative effect following averted gaze, i.e. the difference in RT between imitatively incompatible (slope = 0.85) and compatible trials (slope = -0.23), was significantly stronger with increasing AQ scores, $b = 1.08$, *95% CI* = [0.37, 1.79], $SE = 0.36$, $t = 2.97$, $p = .003$. In other words, individuals with fewer autistic traits seemed to show a lower tendency to imitate the actors with averted gaze than individuals with higher autistic traits. The influence of AQ on imitative compatibility following direct gaze (slope $_{Incomp} = 0.12$; slope $_{Comp} = 0.39$) was not significant, $b = -0.27$, *95% CI* = [-0.97, 0.42], $SE = 0.35$, $t = -0.78$, $p = .44$.

In the preregistered analysis, the AQ modulation of the gaze effects on imitative compatibility did not reach statistical significance (see supplementary Table S3.1). Note, however, that the originally planned approach is not directly comparable to the one reported here. For the preregistered analysis, we first computed imitative compatibility scores for each participant by subtracting the mean RT of imitatively compatible trials from the mean RT of imitatively

incompatible trials. Imitative compatibility scores were then used as the dependent variable in a linear mixed model testing the interaction between AQ and gaze. The preregistered analysis thus examined whether the difference in the level of imitation following direct gaze compared to averted gaze varied across AQ scores, a 2-way interaction that was not statistically significant, b = -0.97, *95% CI* = [-2.36, 0.41], SE = 0.71, t = -1.38, p = .17. In contrast, the model described here tested whether the influence of autistic traits on imitatively compatible and incompatible trials differed for direct gaze and averted gaze conditions.

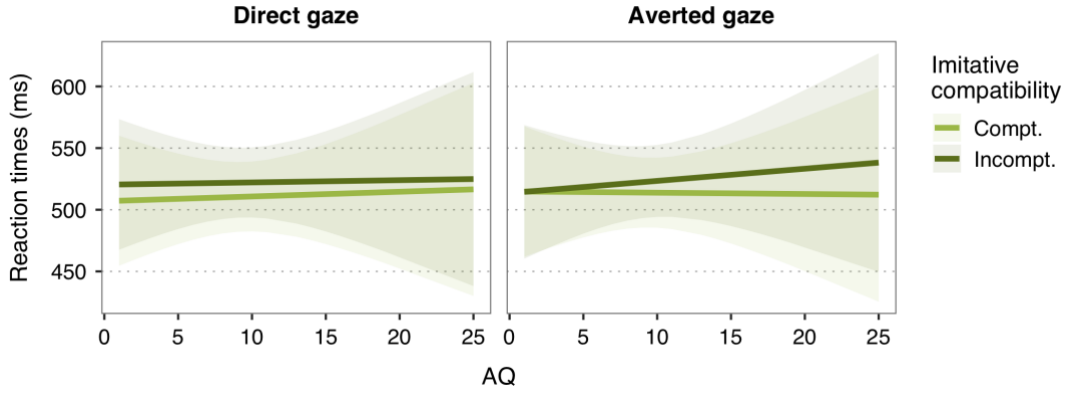

Figure 3.4. Predicted effects of autistic traits (as measured by the Autism Spectrum Quotient; AQ) on reaction times for imitatively compatible (light green line) and imitatively incompatible trials (dark green line), for direct gaze (left panel) and averted gaze (right panel) conditions. For illustration purposes, raw AQ scores are shown in the x-axis, but note that AQ scores were mean-centred before conducting the GLMM. Shaded areas indicate 95% CI.

3.3.1.2 Error rates

The GLMM on error rates yielded significant main effects of imitative compatibility, OR = 1.58, *95% CI* = [1.35, 1.85], SE = 0.08, t = 5.59, p < .001, and spatial compatibility, OR = 1.70, *95% CI* = [1.45, 2.00], SE = 0.08, t = 6.46, p < .001. Compatibility effects indicate that error rates were higher in imitatively incompatible (M = 0.05, SE = 0.004) and spatially incompatible trials (M = 0.05, SE = 0.004) than in imitatively compatible (M = 0.04, SE = 0.003) and spatially compatible trials (M = 0.03, SE = 0.003), respectively (Figure 3.3). None of the remaining main effects or interactions were statistically significant (all p > .07).

3.3.2 Exploratory analyses

3.3.2.1 General compatibility

To make our results comparable to previous automatic imitation studies in which imitative and spatial compatibility were confounded, we ran an additional GLMM on RT data from the subset of trials in which spatial and imitative compatibility overlap (i.e. left-hand trials) (Marsh et al., 2016). This model yielded a main effect of general compatibility, b = 33.11, *95% CI* = [29.45, 36.77], SE = 1.87, t = 17.73, p < .001, with faster correct responses in compatible

(M = 481.00, SE = 1.98) than in incompatible trials (M = 513.15, SE = 117.40). The main effect of gaze was also significant, b = -4.95, *95% CI* = [-8.58, -1.33], SE = 1.85, t = -2.68, p = .007, showing that responses following direct gaze (M = 494.21, SE = 2.00) were faster than following averted gaze (M = 499.10, SE = 2.06). Neither the interaction between gaze and general compatibility, b = -2.65, *95% CI* = [-9.87, 4.58], SE = 3.69, t = -0.72, p = .47, nor the predicted three-way interaction with AQ, b = -1.34, *95% CI* = [-2.68 – 0.01], SE = 0.69, t = -1.95, p = .05, reached the significance threshold. See supplementary Table S3.4 for the full model statistics.

3.3.2.2 Interaction between gaze and the ethnicity of the imitated agent

To explore whether participants reacted differently to the gaze cues of Northern-European actors compared to Mediterranean actors, we ran a GLMM with RT data that included the same predictors as in the confirmatory analysis, except that the variable 'AQ' was replaced by the 2-level categorical factor 'Ethnicity'. The model specification and full statistics are available in the supplementary Table S3.5.

As in all previous models, the main effects of imitative compatibility, spatial compatibility and gaze were significant (all p < .002). In addition, we found a significant 3-way interaction between the actors' ethnicity, gaze and imitative compatibility, b = 13.72, *95% CI* = [4.20, 23.25], SE = 4.86, t = 2.82, p = .005 (Figure 3.5). Follow-up analyses revealed a significant effect of gaze on imitative compatibility for Northern-European actors, b = 9.41, *95% CI* = [2.29, 16.54], SE = 3.64, t = 2.59, p = .01, but not for Mediterranean actors, b = -4.60, *95% CI* = [-11.64, 2.44], SE = 3.59, t = -1.28, p = .20. Specifically, Northern-European identities elicited imitative effects only when they established direct gaze, $M_{\text{Incomp-Compt.}}$= 13.45, SE = 2.60, z = 5.18, p < .001, but not with averted gaze, $M_{\text{Incomp-Compt.}}$= 4.14, SE = 2.60, z = 1.60, p = .11. In contrast, Mediterranean-looking actors triggered imitative compatibility effects in both direct gaze, $M_{\text{Incomp-Compt.}}$= 9.59, SE = 2.60, z = 3.70, p < .001, and averted gaze conditions, $M_{\text{Incomp-Compt.}}$= 14.00, SE = 2.64, z = 5.23, p < .001.

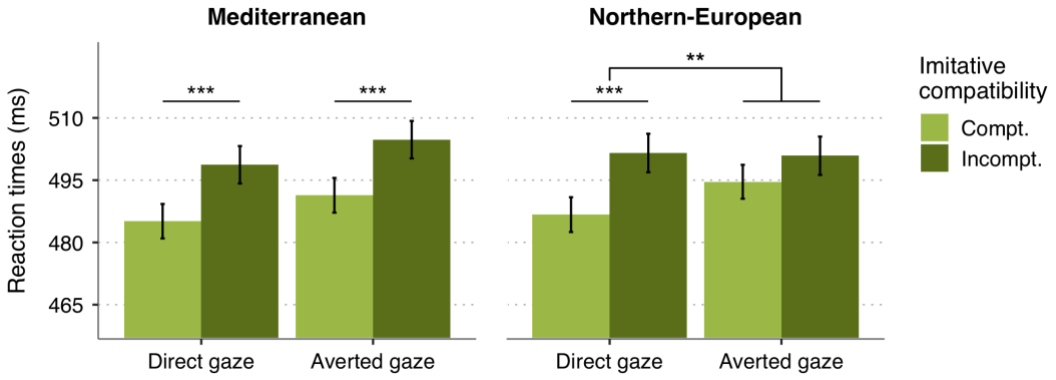

Figure 3.5. Mean reaction times of imitatively compatible (light green bars) and incompatible trials (dark green bars) as a function of gaze cues and ethnicity of the observed actors. Error bars represent within-subject 95% CI. *** p < .001, ** p < .01.

3.3.2.3 Social anxiety

To explore whether symptoms of social anxiety modulated the influence of gaze on automatic imitation, we ran three GLMMs on RT data that included the same predictors as in the confirmatory analyses, except that the variable 'AQ' was replaced by the scores of each of the social anxiety questionnaires (SIAS, SPS and GARS), respectively. The supplementary Table S3.6 includes the specifications and results of these models.

In addition to the significant main effects of imitative compatibility, spatial compatibility and gaze (all $p < .002$), the model with SIAS as predictor revealed a significant 3-way interaction between SIAS, gaze and imitative compatibility, $b = -0.68$, *95% CI* = [-1.09, -0.26], $SE = 0.21$, $t = -3.19$, $p = .001$. Follow-up analyses showed that, as in the case of AQ, the imitative effect in the averted gaze condition was significantly stronger with increasing SIAS scores, $b = 0.37$, *95% CI* = [0.07, 0.67], $SE = 0.15$, $t = 2.40$, $p = .02$. Conversely, imitative effects following direct gaze were weaker as a function of SIAS scores, $b = -0.32$, *95% CI* = [-0.62, -0.02], $SE = 0.15$, $t = -2.12$, $p = .03$. The 3-way interaction between SIAS, gaze and spatial compatibility was not significant, $b = 0.22$, *95% CI* = [-0.20, 0.64], $SE = 0.21$, $t = 1.04$, $p = .30$.

A similar pattern of results was found in the GLMMs with the other two measures of social anxiety: both the 3-way interaction between imitative compatibility, gaze and SPS, $b = -0.72$, *95% CI* = [-1.32, -0.13], $SE = 0.30$, $t = -2.38$, $p = .02$, and between imitative compatibility, gaze and GARS, $b = -0.36$, *95% CI* = [-0.70 – -0.01], $SE = 0.17$, $t = -2.05$, $p = .04$, were significant. No significant interactions were found with spatial compatibility (all $p > 0.35$).

Given that autistic traits and social anxiety seem to modulate the gaze effects on automatic imitation in a similar fashion, and that AQ scores correlated positively with social anxiety measures, all $r_s(60) > .29$, $p < .05$ (Table S3.8), we conducted an additional GLMM to explore whether the observed modulation by AQ could be (partly) attributed to the influence of comorbid social anxiety symptoms. This model included both the 3-way interaction 'AQ*Gaze*Imitative' and 'SIAS*Gaze*Imitative', as well as the corresponding 2-way interactions and main effects for all factors (see supplementary Table S3.7). To avoid overfitting, this model tested the effects on imitative compatibility but not on spatial compatibility. Results showed that, when accounting for the influence of SIAS, the estimated predictive value of the interaction between AQ, gaze and imitative compatibility was lower and no longer statistically significant, $b = -0.56$, *95% CI* = [-1.82, 0.71], $SE = 0.65$, $t = -0.86$, $p = .39$. The 3-way interaction with SIAS remained significant, $b = -0.54$, *95% CI* = [1.08, -0.004], $SE = 0.27$, $t = -1.97$, $p = .048$.

3.3.2.4 Attributed meaning of gaze

Paired-samples *t*-tests were conducted to compare how participants felt in response to direct vs averted gaze clips for each rated attribute. *P*-values were adjusted using the Holm-Bon-

ferroni correction. Participants reported feeling significantly more observed, connected with, accepted, distracted, and pressured by the observed faces following direct gaze compared to averted gaze (all $p < .001$, Cohen's $d > 0.50$). Conversely, averted gaze clips induced more feelings of being ignored and rejected than direct gaze clips (all $p < .001$, Cohen's $d > 0.45$). No significant differences were found between gaze conditions with respect to feeling relieved ($p = 0.31$, Cohen's $d = 0.13$). Summary statistics and complete *t*-tests results are available in the supplementary Table S3.9.

3.4 Discussion

The current preregistered study aimed to resolve inconsistent findings on the influence of gaze cues on automatic imitation. Using a task that disentangled imitative and spatial compatibility, we did not find significant evidence for an overall influence of direct and averted gaze on neither automatic imitation nor spatial processes. However, autistic traits predicted the degree to which participants adapted their imitative behaviour to the gaze cues, such that higher autistic traits were associated with a lower inhibition of imitative responses following averted gaze. In addition, exploratory analyses identified that the ethnicity of the imitated agent, as well as symptoms of social anxiety, may be other modulators of the gaze effects on automatic imitation.

3.4.1 Why did we not replicate the effects of gaze on automatic imitation?

Several factors could explain our failure to replicate the previously reported effect of gaze on automatic imitation. First, participants may have ignored the gaze clips in our study. However, this explanation is unlikely as gaze direction had an overall influence on reaction times, such that participants responded faster following direct gaze than averted gaze. This main effect of gaze has been observed in previous studies (Forbes et al., 2017; Wang, Newort, et al., 2011; Wang & Hamilton, 2014), and is thought to reflect the attention reorienting properties of gaze cues. As shown by eye-tracking data collected in a previous automatic imitation study with gaze clips, direct gaze triggers participants to look at the face region more than averted gaze (Prinsen et al., 2017). In addition, participants in our study reported feeling more observed, connected and accepted when the actors established direct gaze compared to averted gaze, which induced more feelings of rejection and of being ignored. Although these subjective effects were recalled only after the task was finished, they are a further indication that gaze information was processed in accordance with the social evaluations of gaze cues (Wirth et al., 2010).

Second, it could be that gaze alone is not a powerful enough signal to regulate imitative behaviour. In our task, the actors that participants observed maintained a neutral expression throughout the gaze clips. In contrast, the agent presented in previous studies established eye contact with a small smile (Forbes et al., 2017; Prinsen et al., 2017; Wang, Newport, et al.,

2011; Wang & Hamilton, 2014). The combined product of eye contact with a smile, rather than the direct gaze itself, is what may have been interpreted as a signal of social engagement and promoted automatic imitation. In fact, according to the authors, imitative effects during pilot testing in their lab were not elicited if the model established direct gaze without a smile (Wang & Hamilton, 2014). Similarly, even though Marsh et al. (2016) did not find effects on automatic imitation, the influence of gaze on spatial compatibility was conditional upon the group membership of the observed agent. This further suggests that direct and averted gaze may not uniformly influence compatibility effects, but rather the (social) meaning and impact of gaze critically depends on other contextual factors (Hamilton, 2016).

To explore this idea, we examined whether participants in our study reacted differently to the gaze cues depending on the ethnicity of the observed actors. Results showed that the degree of automatic imitation was significantly influenced by the gaze cues of Northern-European actors but not of Mediterranean actors. Specifically, Northern-European identities were only imitated following direct gaze, but not after averted gaze, while Mediterranean actors were automatically imitated regardless of their gaze direction. Although we did not register the ethnicity of our participants, study samples tested in our lab are typically composed of a majority of European descent. Under this assumption, our observations are congruent with gaze cuing studies showing a higher sensitivity to gaze cues of own-race (Collova et al., 2017; Dalmaso et al., 2015; Weisbuch et al., 2017) and high-status faces (Dalmaso et al., 2012; Liuzza et al., 2013). Due to the exploratory nature of this finding, the combined role of gaze and ethnicity on automatic imitation should be taken carefully until further replication. Nevertheless, our results provide tentative evidence that, as speculated above, the influence of gaze on automatic imitation is dependent on other contextual factors.

Our experimental design also deviated from other studies in that eight different identities performed the observed head movements and motor actions. In contrast, most previous paradigms presented the same female actress in all trials (Wang, Newport, et al., 2011; Wang & Hamilton, 2014). Just by the mere-exposure effect, i.e. the increasing preference for a stimulus with repeated exposure (Montoya et al., 2017), participants in earlier studies may have developed increased liking and sympathy for the actress over trials. As familiarity (Deaner et al., 2007) and holding a positive attitude towards the other (Liuzza et al., 2013) seems to enhance sensitivity to gaze cues, increased interest for the observed agent could have made their gazing behaviour more relevant to the participant. This, in turn, may have increased the likelihood to observe gaze effects in previous studies.

Lastly, it could be that our stimuli lacked ecological validity. In a recent TMS study, observation of hand movements following direct gaze elicited stronger cortical motor resonance than after averted gaze in a live two-person context, but not when the gaze cues and hand actions were presented via videotaped recordings (Prinsen & Alaerts, 2019). Though video stimuli used in the current and prior studies allow for better experimental control, they lack

the richness and social meaning inherent in real human interactions, which may hinder the motivation to pay attention to and engage with the stimuli. As such, future studies could further investigate the role of gaze on imitation in more naturalistic settings.

3.4.2 Do autistic traits (and social anxiety) modulate the effects of gaze on automatic imitation?

Previous studies have shown that individuals with ASC tend to automatically imitate others' actions (Cracco, Bardi, et al., 2018; Sowden et al., 2016), but their imitative behaviour is not typically regulated according to the social context (Cook & Bird, 2012; Forbes et al., 2017). In line with this observation, we found that autistic traits in a non-clinical sample predicted the extent to which participants adapted their imitative responses to the gaze cues. However, the modulation by autistic traits reached statistical significance only in the GLMM analysis, not in the preregistered tests. Although this discrepancy could be because mixed-effects models on single-trial data are more powerful than the preregistered analysis plan, results from post-hoc tests further challenge the value of autistic traits in predicting the impact of gaze on automatic imitation.

In particular, exploratory analyses indicated that social anxiety might modulate the effects of gaze on automatic imitation in a similar direction to autistic traits. That is, higher scores in both the AQ (autistic traits) and SIAS (social anxiety) were associated with reduced inhibition of imitative behaviour following averted gaze, and weaker imitative responses following direct gaze, although the latter effect was significant only for SIAS. Given that previous literature typically framed the gaze effects as an enhancement of imitation following direct gaze (Hamilton, 2016; Wang, Newport, et al., 2011; Wang & Hamilton, 2014), the influence of AQ and SIAS on imitative reactions to averted gaze could seem surprising. However, if we consider this type of imitation an automatic process (Heyes, 2011), it is plausible that factors influencing the social modulation of imitation would mostly operate by influencing the ability to inhibit (automatic) imitative responses in situations in which such behaviour would be socially disadvantageous, rather than (or in addition to) enhancing imitation when this would lead to positive social outcomes.

Given the positive correlations between AQ scores and social anxiety measures, the question arose of whether the observed modulation by autistic traits could be related to the concomitant influence of social anxiety. In line with this hypothesis, the predictive value of autistic traits decreased and was no longer statistically significant when the effect of SIAS was accounted for. Because social anxiety symptoms are common in ASC (Spain et al., 2018), future investigations should assess whether the reduced social modulation of imitation previously observed in individuals with ASC (Cook & Bird, 2012; Forbes et al., 2017) could be better explained by comorbid social anxiety. Nevertheless, the stronger weight of SIAS may also be an indication that aspects related to fear of social interactions and atypical social attention,

which are symptoms measured by SIAS and shared with ASC (Hessels et al., 2018; Kliemann et al., 2010; Spain et al., 2018), may be more relevant to the gaze modulation of imitation than other ASC-related constructs assessed by the AQ, such as impairments in communication skills or imagination (Freitag et al., 2007). Future work in larger (sub)clinical samples is needed to disambiguate the patterns of relations between autistic traits and social anxiety underlying the influence of gaze on automatic imitation.

3.4.3 Do social factors modulate automatic imitation independent of spatial compatibility?

Methodological limitations of earlier studies called into question whether social cues thought to regulate automatic imitation could be actually affecting more domain-general mechanisms of response inhibition, such as spatial compatibility (Marsh et al., 2016). In line with this idea, a recent meta-analysis of neuroimaging studies found consistent evidence for the involvement of domain-general brain networks (e.g., dorsolateral frontoparietal cortex) in the control of imitative responses, and only limited support for the engagement of domain-specific systems related to social cognition, such as the theory of mind network (Darda & Ramsey, 2019).

In our study, none of the factors tested (i.e. gaze, ethnicity, autistic traits, social anxiety) showed any significant effects on spatial compatibility. Thus, our data do not support the previous observation that social cues selectively affect spatial compatibility (Marsh et al., 2016). Instead, results from this study are more consistent with the idea that gaze cues, at least in interaction with other factors, modulate automatic imitative responses, even when spatial effects are controlled for. In agreement with the domain-specific hypothesis, neurostimulation studies have implicated the right temporoparietal junction in the control of imitative responses independently of spatial compatibility effects (Sowden & Catmur, 2015), and the medial prefrontal cortex has been identified as a central region in the social modulation of automatic imitation (Wang, Ramsey, et al., 2011; Wang & Hamilton, 2015).

Nevertheless, non-significant results do not prove the absence of an effect. It is possible that social cues also modulate spatial compatibility, but that we were unable to detect it. Because spatial effects are typically stronger in magnitude than imitative effects (here, spatial effects for RT data, b = 22.31, were twice as strong as imitative effects, b = 10.23; Catmur & Heyes, 2011; Marsh et al., 2016), subtle changes by social cues may be more difficult to uncover. The only tentative indication of social modulation of spatial compatibility was a close-to-significant interaction with gaze in the GLMM on RT data, with stronger spatial effects in response to averted gaze compared to direct gaze. However, this effect was in the opposite direction as hypothesized and is not in congruence with the interaction between gaze and group membership reported in Marsh et al. (2016). Altogether, the inconsistent findings with respect to

whether and how social cues impact spatial compatibility call for cautions conclusions regarding the contextual nature of domain-general compatibility effects.

3.4.4 Limitations

Even though our sample size was determined based on the magnitude of the gaze effects reported in previous automatic imitation studies (Forbes et al., 2017; Wang, Newport, et al., 2011), and it included at least twice as many participants, we had limited power to investigate effects of small-to-medium size, which are more likely for interactions with personality traits and contextual factors. Underpowered studies do not only limit the chance of finding an effect but also reduce the likelihood that a statistically significant result reflects a true effect (Button et al., 2013). Therefore, and given the exploratory nature of some of the analyses reported here, further studies with bigger sample sizes are needed to validate the observed social modulation of imitation, as well as the role of autistic traits and social anxiety.

Moreover, this and previous studies on the effects of gaze on imitative behaviour have used tasks that measure imitation of very simple, meaningless finger or hand actions (Forbes et al., 2017; Marsh et al., 2016; Wang, Newport, et al., 2011; Wang & Hamilton, 2014). Though these tasks are widely used in automatic imitation research, they may not fully capture the nature of the actions and gestures that would be spontaneously imitated in real social interactions. Future research could aim to replicate the observed contextual effects of gaze with automatic imitation paradigms that build on more socially meaningful gestures, such as in (Bortoletto et al., 2013; Cracco, Genschow, et al., 2018). Prospective studies would also benefit from including a baseline condition in which no gaze cue precedes the observed irrelevant finger/hand action. A comparison between each gaze condition with the baseline may shed light on whether the gaze effects are due to a stronger tendency to imitate the other's actions following direct gaze and/or reduced imitative responses after averted gaze.

3.4.5 Conclusion

Our study highlights the importance of (preregistered) replications in psychological research. Even though several studies supported the influence of gaze on automatic imitation (Forbes et al., 2017; Prinsen et al., 2017; Wang, Newport, et al., 2011; Wang, Ramsey, et al., 2011; Wang & Hamilton, 2014), most of the successful replication attempts were conducted using the same paradigm as in the original study. Albeit direct replications are crucial to controlling for sampling error, partial and conceptual replications (i.e. experiments that test the same phenomenon with different methodology) are also needed to confirm the internal validity and generalizability of the findings, especially when a research topic appears to be highly sensitive to contextual factors (Schmidt, 2009).

Results from this study indicate that the influence of gaze on automatic imitation may not be as consistent and uniform as reported in the literature. By using different stimuli and task

manipulations, Marsh et al. (2016) and the current study have helped identify potential modulators of the gaze effects on automatic imitation masked before. For example, preliminary findings from our study suggest that characteristic of the imitated agent (e.g., ethnicity), as well as individual differences in social functioning (e.g., autistic traits, social anxiety), should be considered when assessing the role of gaze cues on the regulation of imitative behaviour. Moreover, our results strengthen the idea that imitative and spatial compatibility effects are dissociable processes, although more work is needed to determine how social factors affect them. Future research should systematically test the relationship between the different contextual modulators to better characterize the key elements involved in the social regulation of automatic imitation.

3.5 Supplementary material

3.5.1 Preregistered statistical analyses

Analysis sample

Following the preregistered exclusion criteria, data from 2 participants were removed from these analyses due to the rejection of more than 40% of the trials in one or more conditions. This led to an analysis sample size of 58 participants (30 females, 27 males, 1 non-binary).

Hypothesis 1: Effects of gaze on imitative and spatial compatibility

To test the effects of gaze on imitative and spatial compatibility, we computed the mean reaction times (RT) and mean error rates (ER) for each of the 8 conditions for each participant. Each dependent variable (RT, ER) was analysed with a three-way repeated-measures ANOVA with 'Gaze' (2 levels: direct, averted), 'Spatial compatibility' (2 levels: compatible, incompatible) and 'Imitative compatibility' (2 levels: compatible, incompatible) as within-subject factors.

The ANOVA on RT yielded a significant main effect of both imitative compatibility, $F(1, 57) = 22.04$, $MSE = 727.05$, $p < .001$, $\eta_p^2 = .28$, and spatial compatibility, $F(1, 57) = 49.36$, $MSE = 841.68$, $p < .001$, $\eta_p^2 = .46$. Specifically, participants were faster to perform the correct finger movements when they observed an imitatively compatible action ($M = 494.96$, $SD = 34.46$) compared to an imitatively incompatible action ($M = 506.44$, $SD = 34.92$). Similarly, RT were lower for spatially compatible trials ($M = 491.10$, $SD = 32.65$) compared to incompatible trials ($M = 510.03$, $SD = 33.47$). The main effect of gaze was also statistically significant, $F(1, 57) = 8.55$, $MSE = 307.89$, $p = .005$, $\eta_p^2 = .13$, showing that participants responded faster after direct gaze ($M = 498.18$, $SD = 34.91$) than averted gaze ($M = 502.94$, $SD = 36.11$). Neither of the predicted interactions gaze*imitative compatibility, $F(1, 57) = 2.07$, $MSE = 425.33$, $p = .16$, $\eta_p^2 = .04$, nor gaze*spatial compatibility, $F(1, 57) = 1.57$, $MSE = 320.49$, $p = .21$, $\eta_p^2 = .03$, were statistically significant.

The repeated-measures ANOVA on ER yielded a significant main effect of imitative compatibility, $F(1, 57) = 13.66$, $MSE = 0.002$, $p < .001$, $\eta_p^2 = .19$, such that participants made fewer errors in imitatively compatible trials ($M = 0.03$, $SD = 0.05$) than in incompatible trials ($M = 0.05$, $SD = 0.05$). The main effect of spatial compatibility was also significant, $F(1, 57) = 14.28$, $MSE = 0.001$, $p < .001$, $\eta_p^2 = .20$, with spatially compatible trials ($M = 0.03$, $SD = 0.05$) leading to fewer errors than spatially incompatible trials ($M = 0.04$, $SD = 0.05$). None of the remaining effects reached statistical significance (all $p > .34$).

Hypothesis 2: AQ modulation of the effects of gaze on imitative and spatial compatibility

To test hypothesis 2, we first computed imitative and spatial compatibility scores for each participant and gaze condition, for both RT and ER. Compatibility scores were calculated by subtracting the mean RT [ER] of imitatively [spatially] compatible trials from the mean RT [ER] of imitative [spatially] compatible trials. Positive values indicate the occurrence of a compatibility effect (i.e. slower RT [higher ER] in incompatible trials than compatible trials).

The computed compatibility scores were fitted to linear mixed-effect models (LMM) to test the influence of autistic traits on the effect of gaze on imitative and spatial compatibility. The general model was as follows: DV ~ Gaze + AQ + Gaze*AQ + (1|Participant), where 'Gaze' (categorical predictor with 2 levels: direct, averted), 'AQ' (continuous predictor; mean-centred) and the interaction between the two ('Gaze*AQ') were the fixed effects, and '1|Participant' added random intercepts for participants. Imitative and spatial compatibility scores computed for RT and ER were used as dependent variables. Neither the main effects nor the predicted interaction 'Gaze*AQ' reached statistical significance in any of the four models tested (all $p > .15$; Table S3.1).

Table S3.1. Results of the linear mixed models conducted according to the preregistered analyses.

	Imitative compatibility scores								Spatial compatibility scores							
	Model 1: Reaction times				Model 2: Error rates				Model 3: Reaction times				Model 4: Error rates			
	b *(95% CI)*	*SE*	*t*	*p*	*Odd Ratios* *(95% CI)*	*SE*	*t*	*p*	*b* *(95% CI)*	*SE*	*t*	*p*	*Odd Ratios* *(95% CI)*	*SE*	*t*	*p*
Gaze	5.51 (-1.94 – 12.96)	3.80	1.45	0.15	0.001 (-0.010 – 0.012)	0.006	0.24	0.81	-4.17 (-10.70 – 2.35)	3.33	-1.25	0.21	0.006 (-0.007 – 0.019)	0.006	0.95	0.34
AQ	0.47 (-0.44 – 1.39)	0.47	1.02	0.31	0.001 (-0.001 – 0.002)	0.001	0.92	0.36	-0.24 (-1.23 – 0.74)	0.50	-0.49	0.63	-0.0004 (-0.002 – 0.001)	0.001	-0.64	0.52
Gaze*AQ	-0.97 (-2.36 – 0.41)	0.71	-1.38	0.17	0.0002 (-0.002 – 0.002)	0.001	0.18	0.86	-0.56 (-1.78 – 0.65)	0.62	-0.91	0.36	-0.001 (-0.003 – 0.001)	0.001	-0.73	0.47
Random Effects																
σ2	418.72				0.001				321.50				0.002			
τ00	153.91 $_{\text{Participant}}$				0.001 $_{\text{Participant}}$				265.81 $_{\text{Participant}}$				0.001 $_{\text{Participant}}$			
ICC	0.27				0.43				0.45				0.09			
N	58 $_{\text{Participant}}$				58 $_{\text{Participant}}$				58 $_{\text{Participant}}$				58 $_{\text{Participant}}$			
Observations	116				116				116				116			
Marginal R2	0.036				0.011				0.014				0.015			
Conditional R2	0.295				0.437				0.460				0.103			

Note. p-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Model equation: Compt. scores ~ Gaze + AQ + AQ*Gaze + (1|Participant) + (1|Stimulus). AQ: Autism Spectrum Quotient; b: unstandardized coefficient; SE: standard error; ß: standardized coefficient; CI: confidence interval; t: test statistic coefficient; p: p-value; σ2: within-group variance; τ00 = between-group variance; ICC = interclass correlation (ratio of between-cluster variance to total variance); N: number of random effects.

3.5.2 Supplementary tables: Confirmatory analyses

Table S3.2. Mean and standard deviation of reaction times (ms) and error rates for each imitative and spatial compatibility condition, as a function of gaze.

		Reaction times				Error rates			
		Direct gaze		Averted gaze		Direct gaze		Averted gaze	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Imitative compatibility	Compatible	485.91	96.79	492.97	95.82	0.04	0.21	0.03	0.21
	Incompatible	500.12	105.14	502.80	104.59	0.05	0.25	0.05	0.25
Spatial compatibility	Compatible	484.32	105.50	486.82	104.34	0.03	0.20	0.04	0.21
	Incompatible	501.91	95.76	509.14	94.43	0.06	0.26	0.05	0.25

Table S3.3. Results of the generalized linear mixed-effects models conducted for the confirmatory analyses.

	Model 1: Reaction times						Model 2: Error rates					
	b	*95% CI*	*SE*	*ß*	*t*	*p*	*Odds Ratios*	*95% CI*	*SE*	*ß*	*t*	*p*
Imitative compt.	10.23	7.57 – 12.90	1.36	0.05	7.53	<0.001	1.58	1.35 – 1.85	0.08	1.12	5.59	<0.001
Spatial compt.	22.31	19.63 – 24.99	1.37	0.10	16.34	<0.001	1.70	1.45 – 2.00	0.08	1.30	6.46	<0.001
Gaze	-4.62	-7.29 – -1.96	1.36	-0.02	-3.41	0.001	0.99	0.84 – 1.16	0.08	-0.03	-0.15	0.877
AQ	0.29	-4.97 – 5.54	2.68	0.01	0.11	0.915	0.97	0.92 – 1.02	0.03	-0.75	-1.09	0.274
Gaze * Imitative	2.51	-2.75 – 7.78	2.69	0.01	0.93	0.350	0.94	0.68 – 1.29	0.16	-0.08	-0.39	0.698
Gaze * Spatial	-5.10	-10.44 – 0.23	2.72	-0.01	-1.87	0.061	1.12	0.81 – 1.54	0.16	0.14	0.67	0.503
AQ * Imitative	0.45	-0.04 – 0.94	0.25	0.01	1.79	0.074	1.03	1.00 – 1.06	0.02	0.36	1.79	0.073
AQ * Spatial	-0.15	-0.64 – 0.34	0.25	-0.004	-0.59	0.552	1.01	0.98 – 1.04	0.02	0.10	0.50	0.614
AQ * Gaze	-0.08	-0.57 – 0.41	0.25	-0.002	-0.31	0.755	1.01	0.98 – 1.04	0.02	0.14	0.69	0.490
AQ * Gaze * Imitative	-1.29	-2.28 – -0.31	0.50	-0.02	-2.59	0.010	1.00	0.94 – 1.06	0.03	-0.01	-0.06	0.953
AQ * Gaze * Spatial	0.16	-0.83 – 1.14	0.50	0.002	0.31	0.754	0.97	0.91 – 1.03	0.03	-0.21	-1.05	0.294
Random Effects												
σ^2	0.0001						3.29					
τ_{00}	$32.25_{Stimulus}$, $620.88_{Participant}$						$0.03_{Stimulus}$, $1.00_{Participant}$					
ICC	-						0.24					
N	$60_{Participant}$, $64_{Stimulus}$						$60_{Participant}$, $64_{Stimulus}$					
Observations	14047						15353					
Marginal R^2/ Conditional R^2	-						0.035 / 0.265					

Note. p-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Inter-class correlation (ICC) and R^2 coefficients are not reported for Model 1 as available methods for the estimation of such parameters are not optimized for GLMM with inverse-Gaussian distributions (Nakagawa et al., 2017). Model equation: DV ~ Imitative + Spatial + Gaze + AQ + Gaze*Imitative + Gaze*Spatial + AQ*Imitative + AQ*Spatial + AQ*Gaze + AQ*Gaze*Imitative + AQ*Gaze*Spatial + (1|Participant) + (1|Stimulus). AQ: Autism Spectrum Quotient; *b*: unstandardized coefficient; *SE*: standard error; *ß*: standardized coefficient; *CI*: confidence interval; *t*: test statistic coefficient; *p*: p-value; σ2: within-group variance; τ00 = between-group variance; ICC = interclass correlation (ratio of between-cluster variance to total variance); N: number of random effects.

3.5.3 Supplementary tables: Exploratory analyses

Table S3.4. Results of the generalized linear mixed-effects model used to explore the influence of gaze and autistic traits on general compatibility.

	Reaction times					
	b	*95% CI*	*SE*	*ß*	*t*	*p*
General compt.	33.11	29.45 – 36.77	1.87	0.15	17.73	<0.001
Gaze	-4.95	-8.58 – -1.33	1.85	-0.02	-2.68	0.007
AQ	0.11	-5.27 – 5.49	2.75	0.01	0.04	0.967
Gaze * General compt.	-2.65	-9.87 – 4.58	3.69	-0.01	-0.72	0.473
AQ * General compt.	0.22	-0.45 – 0.89	0.34	0.01	0.64	0.524
AQ * Gaze	-0.56	-1.24 – 0.11	0.34	-0.01	-1.65	0.100
AQ * Gaze * General compt.	-1.34	-2.68 – 0.01	0.69	-0.02	-1.95	0.051
Random Effects						
σ2	0.0001					
τ00	50.81 $_{\text{Stimulus}}$, 857.63 $_{\text{Participant}}$					
N	60 $_{\text{Participant}}$, 64 $_{\text{Stimulus}}$					
Observations	7010					

Note. *p*-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Inter-class correlation (ICC) and R^2 coefficients are not reported as available methods for the estimation of such parameters are not optimized for GLMM with inverse-Gaussian distributions (Nakagawa, Johnson & Schielzeth, 2017). Model equation: RT ~ General compt. + Gaze + AQ + Gaze*General compt. + AQ*General compt. + AQ*Gaze + AQ*Gaze*General compt. + (1|Participant) + (1|Stimulus). AQ: Autism Spectrum Quotient; *b*: unstandardized coefficient; *SE*: standard error; *ß*: standardized coefficient; *CI*: confidence interval; *t*: test statistic coefficient; *p*: p-value; σ2: within-group variance; τ00 = between-group variance; ICC = interclass correlation (ratio of between-cluster variance to total variance); N: number of random effects.

Table S3.5. Results of the generalized linear mixed-effects model used to explore the influence of ethnicity and gaze on imitative and spatial compatibility.

	Reaction times					
	b	*95% CI*	*SE*	*ß*	*t*	*p*
Imitative compt.	10.30	7.64 – 12.95	1.36	0.05	7.60	<0.001
Spatial compt.	22.29	19.65 – 24.92	1.34	0.10	16.57	<0.001
Gaze	-4.65	-7.30 – -2.00	1.35	-0.02	-3.44	0.001
Ethnicity	1.02	-4.22 – 6.26	2.67	0.01	0.38	0.702
Gaze * Imitative	2.45	-2.68 – 7.59	2.62	0.01	0.94	0.350
Gaze * Spatial	-5.06	-10.20 – 0.08	2.62	-0.01	-1.93	0.054
Ethnicity * Imitative	-3.00	-8.22 – 2.22	2.66	-0.01	-1.13	0.260
Ethnicity * Spatial	-4.18	-9.42 – 1.07	2.68	-0.01	-1.56	0.118
Ethnicity * Gaze	4.01	-1.23 – 9.26	2.68	0.01	1.50	0.134
Ethnicity * Gaze * Imitative	13.72	4.20 – 23.25	4.86	0.02	2.82	0.005
Ethnicity * Gaze * Spatial	-3.54	-12.74 – 5.66	4.69	-0.004	-0.75	0.451
Random effects						
σ2	0.0001					
τ00	32.13 $_{\text{Stimulus}}$, 621.05 $_{\text{Participant}}$					
N	60 $_{\text{Participant}}$, 64 $_{\text{Stimulus}}$					
Observations	14047					

Note. *p*-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Inter-class correlation (ICC) and R^2 coefficients are not reported as available methods for the estimation of such parameters are not optimized for GLMM with inverse-Gaussian distributions (Nakagawa, Johnson & Schielzeth, 2017). Model equation: RT ~ Imitative + Spatial + Gaze + Ethnicity + Gaze*Imitative + Gaze*Spatial + Ethnicity*Imitative + Ethnicity*Spatial + Ethnicity*Gaze + Ethnicity*Gaze*Imitative + Ethnicity*Gaze*Spatial + (1|Participant) + (1|Stimulus). *b*: unstandardized coefficient; *SE*: standard error; *ß*: standardized coefficient; *CI*: confidence interval; *t*: test statistic coefficient; *p*: p-value; σ^2: within-group variance; τ_{00} = between-group variance; N: number of random effects.

Table S3.6. Results of the generalized linear mixed-effects models used to explore the modulation of social anxiety traits on the influence of gaze on imitative and spatial compatibility.

	Model 1: SIAS						Model 2: SPS						Model 3: GARS					
	b	*95% CI*	*SE*	*ß*	*t*	*p*	*b*	*95% CI*	*SE*	*ß*	*t*	*p*	*b*	*95% CI*	*SE*	*ß*	*t*	*p*
Imitative compt.	10.26	7.60 – 12.92	1.36	0.05	7.56	<0.001	10.16	7.49 – 12.83	1.36	0.05	7.46	<0.001	10.29	7.64 – 12.94	1.35	0.05	7.61	<0.001
Spatial compt.	22.20	19.54 – 24.86	1.36	0.10	16.36	<0.001	22.21	19.53 – 24.88	1.36	0.10	16.30	<0.001	22.28	19.64 – 24.91	1.34	0.10	16.58	<0.001
Gaze	-4.57	-7.23 – -1.90	1.36	-0.02	-3.36	0.001	-4.47	-7.15 – -1.80	1.36	-0.02	-3.28	0.001	-4.62	-7.27 – -1.97	1.35	-0.02	-3.42	0.001
Quest.	-0.69	-3.27 – 1.88	1.31	-0.08	-0.53	0.598	-1.21	-4.83 – 2.40	1.84	-0.09	-0.66	0.511	0.34	-1.76 – 2.44	1.07	0.05	0.32	0.750
Gaze * Imitative	2.97	-2.36 – 8.29	2.72	0.01	1.09	0.275	2.91	-2.33 – 8.15	2.67	0.01	1.09	0.277	2.50	-2.66 – 7.67	2.63	0.01	0.95	0.342
Gaze * Spatial	-5.34	-10.61 – -0.07	2.69	-0.01	-1.99	0.047	-5.15	-10.48 – 0.19	2.72	-0.01	-1.89	0.059	-5.13	-10.40 – 0.14	2.69	-0.01	-1.91	0.057
Quest.* Imitative	0.05	-0.16 – 0.26	0.11	0.003	0.47	0.639	0.18	-0.12 – 0.48	0.15	0.01	1.17	0.241	-0.04	-0.21 – 0.13	0.09	-0.003	-0.45	0.653
Quest.* Spatial	0.12	-0.08 – 0.33	0.11	0.01	1.16	0.244	0.17	-0.13 – 0.47	0.15	0.01	1.14	0.255	0.05	-0.12 – 0.22	0.09	0.003	0.56	0.576
Quest.* Gaze	-0.08	-0.29 – 0.13	0.11	-0.004	-0.78	0.435	-0.23	-0.53 – 0.07	0.15	-0.01	-1.50	0.134	-0.06	-0.23 – 0.11	0.09	-0.004	-0.72	0.471
Quest.* Gaze * Imitative	-0.68	-1.09 – -0.26	0.21	-0.02	-3.19	0.001	-0.72	-1.32 – -0.13	0.30	-0.01	-2.38	0.017	-0.36	-0.70 – -0.01	0.17	-0.01	-2.05	0.041
Quest.* Gaze * Spatial	0.22	-0.20 – 0.64	0.21	0.01	1.04	0.299	0.02	-0.58 – 0.62	0.31	0.0001	0.07	0.947	0.15	-0.19 – 0.49	0.17	0.01	0.88	0.379
Random Effects																		
σ2	0.0001						0.0001						0.0001					
τ00	31.80_{Stimulus}, $651.78_{\text{Participant}}$						31.79_{Stimulus}, $649.58_{\text{Participant}}$						31.63_{Stimulus}, $655.59_{\text{Participant}}$					
N	$60_{\text{Participant}}$, 64_{Stimulus}						$60_{\text{Participant}}$, 64_{Stimulus}						$60_{\text{Participant}}$, 64_{Stimulus}					
Observations	14047						14047						14047					

Note. The predictor "Quest." refers to the corresponding questionnaire of social anxiety (SIAS, SPS or GARS). *p*-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Inter-class correlation (ICC) and R^2 coefficients are not reported as available methods for the estimation of such parameters are not optimized for GLMM with inverse-Gaussian distributions (Nakagawa, Johnson & Schielzeth, 2017). Model equation: RT ~ Imitative + Spatial + Gaze + Questionnaire + Gaze*Imitative + Gaze*Spatial + Questionnaire*Imitative + Questionnaire*Spatial + Questionnaire*Gaze + Questionnaire*Gaze*Imitative + Questionnaire *Gaze*Spatial + (1|Participant) + (1|Stimulus). SIAS: Social Interaction Anxiety Scale; SPS: Social Phobia Scale; GARS: Gaze Anxiety Rating Scale; *b*: unstandardized coefficient; *SE*: standard error; *ß*: standardized coefficient; *CI*: confidence interval; *t*: test statistic coefficient; *p*: p-value; σ^2: within-group variance; τ_{00} = between-group variance; N: number of random effects.

Table S3.7. Results of the generalized linear mixed-effects model used to explore the modulation of AQ and SIAS on the effect of gaze on imitative compatibility.

	Reaction times					
	b	*95% CI*	*SE*	*ß*	*t*	*p*
Imitative compt.	10.44	7.75 – 13.14	1.37	0.05	7.60	<0.001
Gaze	-4.61	-7.30 – -1.92	1.37	-0.02	-3.36	0.001
SIAS	-1.38	-4.72 – 1.97	1.71	-0.15	-0.81	0.420
AQ	2.23	-4.99 – 9.46	3.69	0.11	0.61	0.545
Gaze * Imitative	2.82	-2.47 – 8.11	2.70	0.01	1.04	0.296
SIAS * Imitative	-0.14	-0.41 – 0.13	0.14	-0.01	-1.00	0.318
SIAS * Gaze	-0.10	-0.37 – 0.17	0.14	-0.01	-0.71	0.480
AQ * Imitative	0.60	-0.04 – 1.24	0.33	0.01	1.85	0.065
AQ * Gaze	0.09	-0.55 – 0.72	0.32	0.002	0.27	0.786
SIAS * Gaze * Imitative	-0.54	-1.08 – -0.00	0.27	-0.02	-1.97	0.048
AQ * Gaze * Imitative	-0.56	-1.82 – 0.71	0.65	-0.01	-0.86	0.388
Random effects						
σ2	0.0001					
τ00	30.82 $_{\text{Stimulus}}$, 641.04$_{\text{Participant}}$					
N	60 $_{\text{Participant}}$, 64 $_{\text{Stimulus}}$					
Observations	14047					

Note. *p*-values for the fixed effects calculated using Wald-statistics approximation, uncorrected. Inter-class correlation (ICC) and R^2 coefficients are not reported as available methods for the estimation of such parameters are not optimized for GLMM with inverse-Gaussian distributions (Nakagawa, Johnson & Schielzeth, 2017). Model equation: RT ~ Imitative + Gaze + SIAS + AQ + Gaze*Imitative + SIAS*Imitative + SIAS*Gaze + AQ*Imitative + AQ*Gaze + SIAS*Gaze*Imitative + AQ*Gaze*Spatial + (1|Participant) + (1|Stimulus). *b*: unstandardized coefficient; *SE*: standard error; *ß*: standardized coefficient; *CI*: confidence interval; *t*: test statistic coefficient; *p*: p-value; σ^2: within-group variance; τ_{00} = between-group variance; N: number of random effects.

Table S3.8. Descriptive statistics, internal consistency reliability, and correlations between questionnaires of autistic traits and social anxiety.

	M	*SD*	Range	α	1	2	3
1. AQ	9.47	5.41	1 – 25	.82			
2. SIAS	26.43	12.15	4 – 59	.89	.64***		
					[.46, .77]		
3. SPS	12.20	8.67	2 – 43	.88	.44***	.68***	
					[.21, .64]	[.49, .81]	
4. GARS	22.72	14.87	0 – 67	.94	.29*	.59***	.50***
					[.04, .52]	[.39, .74]	[.30, .67]

Note. Cronbach's alpha was used as an index of internal consistency. Correlation coefficients were computed with Spearman's rank correlations. Values in square brackets indicate the 95% confidence interval for each correlation. M: mean; SD: standard deviation; α: Cronbach's alpha; AQ: Autism Spectrum Quotient (33-items version; Freitag et al., 2007); SIAS: Social Interaction Anxiety Scale (Stangier et al., 1999); SPS: Social Phobia Scale (Stangier et al., 1999); GARS: Gaze Anxiety Rating Scale (Domes et al., 2016); * $p < .05$; ** $p < .01$; *** $p < .001$.

Table S3.9. Summary and t-test statistics for the ratings on the attributed meaning of direct and averted gaze.

	Mean (*SD*)		*T*-test	Cohen's *d*
	Direct gaze	Averted gaze		
Observed	1.73 (1.29)	0.30 (0.65)	$t(59) = 8.33$, 95% *CI* [1.09, 1.78], $p < .001$	1.08
Connected	1.55 (1.23)	0.57 (0.77)	$t(59) = 7.15$, 95% *CI* [0.71, 1.26], $p < .001$	0.92
Accepted	1.30 (1.14)	0.45 (0.70)	$t(59) = 6.24$, 95% *CI* [0.58, 1.12], $p < .001$	0.81
Distracted	2.12 (1.12)	1.28 (1.03)	$t(59) = 5.05$, 95% *CI* [0.50, 1.16], $p < .001$	0.65
Pressured	1.32 (1.11)	0.67 (0.95)	$t(59) = 3.95$, 95% *CI* [0.32, 0.98], $p = .001$	0.51
Ignored	0.28 (0.67)	1.27 (1.18)	$t(59) = -6.27$, 95% *CI* [-1.30, -0.67], $p < .001$	0.81
Rejected	0.52 (0.77)	1.03 (0.94)	$t(59) = -3.60$, 95% *CI* [-0.80, -0.23], $p = .001$	0.46
Relieved	0.93 (1.09)	1.12 (1.11)	$t(59) = -1.02$, 95% *CI* [-0.54, 0.18], $p = .31$	0.13

Note. P-values were adjusted based on the Holm-Bonferroni method to correct for multiple comparisons.

4. Study 3: Emotional egocentricity & socio-emotional traits

Affective states influence emotion perception: evidence for emotional egocentricity

Irene Trilla, Anne Weigand and Isabel Dziobek

Abstract. Research in social cognition has shown that our own emotional experiences are an important source of information to understand what other people are feeling. The current study investigated whether individuals project their own affective states when reading other's emotional expressions. We used brief autobiographical recall and audio-visual stimuli to induce happy, neutral and sad transient states. After each emotion induction, participants made emotion judgments about ambiguous faces displaying a mixture of happiness and sadness. Using an adaptive psychophysics procedure, we estimated the tendency to perceive the faces as happy under each induced affective states. Results demonstrate the occurrence of egocentric projections, such that faces were more likely judged as happy when participants reported being happy compared to when they were sad. Moreover, the degree of emotional egocentricity was associated with individual differences in perspective-taking, with smaller biases being observed in individuals with a higher disposition to take others' perspective. Our findings extend previous literature on emotional egocentricity by showing that self-projection occurs when we make emotion attributions based on the other's emotional expressions, and supports the notion that perspective-taking tendencies play a role in the ability to understand the other's affective states.

This article has been published as:

Trilla, I., Weigand, A., & Dziobek, I. (2020). Affective states influence emotion perception: Evidence for emotional egocentricity. *Psychological Research.*
https://doi.org/10.1007/s00426-020-01314-3

4.1 Introduction

The ability to understand the emotions of others is crucial to successfully function in social interactions. Due to a lack of direct access to other people's mind, we have to rely on indirect sources of information to infer how others feel. For example, we could use information about the context the other person is exposed to or read their emotional expressions. While these external cues about the other can be valuable means to emotion understanding, research in social cognition has shown that one's own affective state is also used to guide judgments about others' emotional experiences (Silani et al., 2013; Steinbeis & Singer, 2014). That is, people tend to project their own emotions when inferring what other people feel, a process known as emotional egocentricity. Self-projection can be an efficient heuristic, especially when our experiences are similar to the other's. However, it can also lead to inaccurate emotion attributions unless egocentric inferences are adjusted to account for dissimilarities between oneself and the other person (Mitchell, 2009).

Emotional egocentricity has been studied using tasks in which participants are asked to make emotion judgments about themselves and another person while being simultaneously exposed to either affectively congruent stimulation (e.g., both were touched by a pleasant material) or affectively incongruent stimulation (e.g., the participant received pleasant touch, while the target was touched by an unpleasant material; Silani et al., 2013). A tendency to project one's own emotions onto others is typically indicated by emotion judgments biased towards the participants' own affective states, particularly in incongruent conditions. Egocentric biases have consistently been observed in similar perspective-taking paradigms that used monetary reward and punishment (Steinbeis & Singer, 2014), as well as visuo-gustatory (Hoffmann et al., 2015) and audio-visual (von Mohr et al., 2019) stimulation to induce congruent and incongruent affective states to the participant and the target.

A common feature of the existing emotional egocentricity tasks is that participants are not able to see the target's reactions to the affective stimulations, so emotion judgments are exclusively based on information about the type of stimulation the other is exposed to. Under these conditions, social cognitive processes such as perspective-taking may be activated to infer the other's emotion. In particular, performance in these paradigms has been taken as an indicator of the participants' self-other distinction abilities, as egocentric biases in this context are thought to reflect a failure to distinguish the representation of one's own affective states from that of the other (Silani et al., 2013; Hoffmann, Koehne, et al., 2016; Tomova et al., 2014).

In daily life situations, however, we can often rely on more basic abilities that do not require perspective-taking processes to understand what others are feeling, such as emotion perception. Imagine, for example, that you give a present to a friend. You will probably first judge whether they liked it or not based on interpreting their emotional reactions when they un-

wrap it. Indeed, the accurate reading of emotional signals such as facial expressions (Lindner & Rosén, 2006), body postures (de Gelder et al., 2015) or speech prosody (Golan et al., 2007) has been shown to be key for understanding the affective states of others. Until now, however, little attention has been placed on studying egocentricity during perception-based emotion attribution.

Previous research on the influence of mood on emotion perception provides a first indication of egocentric biases when reading others' emotional states. Studies inducing positive and negative affective states to participants have shown that emotional facial expressions are more easily recognized when they are congruent with the participant's induced mood (Lee et al., 2008; Niedenthal et al., 2000; Qiao-Tasserit et al., 2017; Schmid & Schmid Mast, 2010). These mood-congruency effects have been often contextualized under general cognitive theories of affect congruence, according to which affective states activate linked memory representations and facilitate the encoding and processing of affectively-congruent information (Forgas, 2017). However, mood-congruency effects could also reflect emotional egocentricity: they may result from an over-attribution of one's own affective states to others. In line with this interpretation, biases in emotion perception seem to be stronger when the participant's emotion and the emotion expressed by the target are incongruent (Schmid & Schmid Mast, 2010). Compared to classic emotional egocentricity paradigms, however, these biases may stem from more implicit and unconscious processes of self-projection, rather than reflecting self-other distinction abilities.

The current study sought to revisit mood-congruent biases in emotion perception as a measure of emotional egocentricity. First, we developed a novel approach to estimate the degree to which one's own affective states bias judgments of emotional facial expressions. Using a combination of brief autobiographical recall and audio-visual stimuli, we induced happy, neutral and sad transient states to the participants. After each emotion induction, participants completed a short emotion perception task in which they made binary decisions ("happy" or "sad"?) about the expression of faces displaying a mixture of happiness and sadness. We hypothesized that emotion judgments would be biased by the participants' affective states, such that they would more likely judge the ambiguous faces as happy when feeling happy than when being sad.

Second, we predicted that the magnitude of egocentric biases during the emotion judgments would be related to the participants' disposition to consider and react to other people's experiences. In particular, we examined associations with two components of dispositional empathy measured with the Interpersonal Reactivity Index (IRI; Davis, 1980). On the one hand, the empathic concern scale taps into affective empathy and measures the tendency to react with feelings of sympathy and concern for unfortunate others (Davis, 1980). On the other hand, the perspective-taking scale assesses the tendency to adopt the point of view of another person, a facet of cognitive empathy (Davis, 1980). An association between mood-

congruent biases and dispositional empathy would indicate that the mood effects on emotion perception are related to processes of social cognition. Finally, we explored associations with autistic traits, as stronger egocentricity during cognitive mentalizing (Bradford et al., 2018; Pearson et al., 2013) and deficits in emotion recognition (Uljarevic & Hamilton, 2013) are commonly observed in autism spectrum conditions (ASC).

4.2 Methods

4.2.1 Participants

Fifty German-speaking adults (31 females, M_{age} = 27.82, SD_{age} = 6.66, $range_{age}$ = 19 - 44) were recruited for this study. An a-priori power analyses using G*Power 3 (Faul et al., 2007) estimated a sample of 49 participants (alpha = 0.05, power = 0.80, three measurement levels, within-subject repeated-measures analysis of variance) for a η_p^2 = 0.17. The effect size was determined based on the emotional egocentricity effects reported in Silani et al. (2013; behavioural experiment 1: η_p^2 = 0.074; behavioural experiment 2: η_p^2 = 0.277) and Hoffmann et al. (2016; η_p^2 = 0.180).

Exclusion criteria included current psychiatric or neurological disorders, cognitive or neurological impairments, and psychoactive medication. All participants gave written informed consent and were financially remunerated for their participation. The study was conducted in compliance with the latest revision of the Code of Ethics of the World Medical Association (Declaration of Helsinki) and was approved by the Ethics Committee of the Psychology department at Humboldt-Universität zu Berlin.

4.2.2 Materials and procedure

4.2.2.1 Emotional egocentricity paradigm

The emotional egocentricity paradigm comprised three blocks, corresponding to the three affective state manipulations applied to each participant (happy, neutral and sad). Each block was divided into two parts: an emotion induction, followed by an emotion perception task. MATLAB R2015b (The MathWorks, Inc., Natick, Massachusetts, United States) and the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007) were used for stimulus presentation.

Emotion induction. A combination of a brief autobiographical recall and audio-visual clips was used to induce transient happy, neutral and sad states. At the beginning of each block, participants were asked to remember an event in their lives that elicited one of the two target emotions, or a neutral state. Participants wrote down keywords that reminded them of that particular event and were given 4 minutes to imagine themselves in that situation and to relive the emotions they felt at that time. To elicit happiness, participants were asked to

think of an enjoyable moment with friends or children, or a time in which they cuddled or fooled around with a pet. To elicit sadness, participants were prompted to think of a person suffering or the death of a loved one. For the neutral block, participants had to recall a morning routine. The content of these memories was selected to match the theme of the 60-seconds clips they would watch immediately after.

Each clip included 11 pictures presented for 4.5 seconds, with 1-second cross-dissolve transitions in between. Images for the clips were taken from the International Affective Picture System (IAPS; Lang et al., 2005), the Nencki Affective Picture System (NAPS; Marchewka et al., 2013; Riegel et al., 2016) or collected by the experimenters. All emotional pictures had been previously rated by 22 participants in a pilot study as respectively evoking sadness (M = 4.03, SE = 0.53) and happiness (M = 8.50, SE = 0.31) on a 11-point scale (1 = sad, 11 = happy). Examples of sad-inducing pictures are crying children or extreme poverty scenes, while happy-inducing images included smiling children, baby animals, or people dancing. Images used in the neutral clip depicted daily life objects such as tableware, and were selected from the IAPS and NAPS databases (valence ratings between 4.50 and 5.50; both databases used a 9-point scale). The presentation of pictures was accompanied by audio matching in valence. An excerpt from American Honey by Sam & Chesney was used for the happy clip; Adagio for Strings by Samuel Barber was used for the sad clips; and kitchen sounds were used for the neutral clip.

Emotion perception task. A forced-choice psychophysical procedure was used to assess the perception of ambiguous emotional faces after each emotion induction. Twenty-one morphs of a female face were generated by mixing a happy facial expression with a sad facial expression in steps of 5%. The original stimuli were selected from the FACES database (Ebner et al., 2010). Face morphs (496 x 659 pixels) were grey-scaled and embedded within a grey oval that occluded the hair and clothing (Figure 4.1).

Each trial of the task started with a fixation cross shown for 1500 ms, followed by a face morph for 500 ms. Next, the response options ("happy", "sad") appeared on the right and left side of the screen, respectively, until the participant made their response via key press (Figure 4.1). A 1-up/1-down adaptive method (Leek, 2001) was used to estimate the point on the happy-sad morph continuum at which observers were equally likely to judge the target emotion as happy or sad (point of subjective equality; PSE). PSEs have been used in previous studies to measure emotion perception biases (Harris et al., 2016; Marneweck et al., 2013). Here, lower PSEs indicate a higher tendency to perceive faces as happy (i.e. less happiness is required in the face to be equally likely judged as happy or sad). The adaptive procedure was implemented in MATLAB using the Palamedes Toolbox (Prins & Kingdom, 2018). One ascending and one descending interleaved staircases were set to select the morph stimulus on a trial-by-trial basis based on the participant's previous response. Specifically, the morph level presented in each subsequent trial within a staircase was increased one step in the

happy-sad morph continuum (i.e. 5% happier face) if the previous trial was a "sad" response, and decreased one step (i.e. 5% sadder face) if the previous trial was a "happy" response. The descending staircase started with a 90-10% happy-sad morph, and the ascending with a 30-70% happy-sad morph. The starting morph level for each staircase was determined based on pilot data, which showed that PSE with our stimuli was around the 60-40% happy-sad morph. Each staircase stopped after eight reversals. A reversal is defined as a transition point within a staircase in which the participant switched their response from perceiving a happy face to a sad face, or vice versa. The PSE was estimated by averaging the morph levels of the last five reversals of both staircases. On average, participants completed 37.96 trials per block (*SD* = 5.83), which corresponds to a task length of approximately 1.5 minutes. A short task duration was important to ensure that PSEs were measured within the duration of the emotion induction effects.

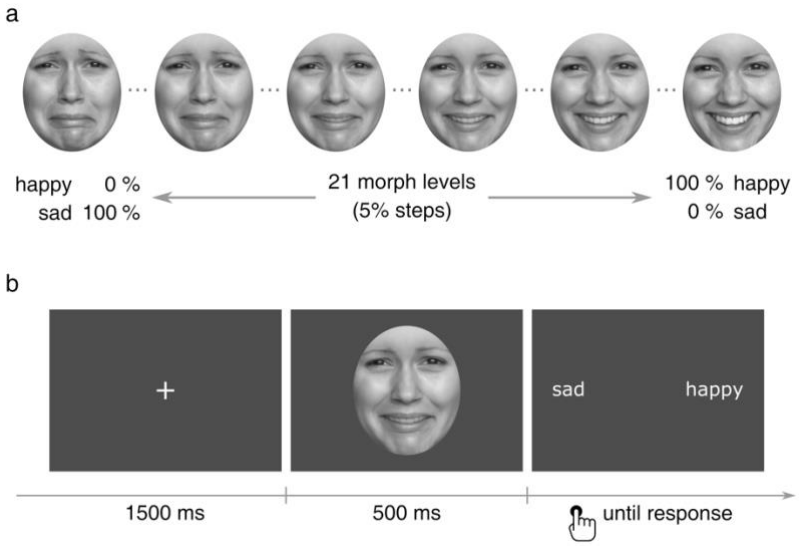

Figure 4.1. (a) Example of the continuum of sad-happy face morphs used in the emotion perception task. Due to copyright restrictions, the depicted face identity does not correspond to the one used in the study. (b) Example of a trial of the emotion perception task. The morph level presented was selected on a trial-by-trial basis using a 1-up/1-down adaptive procedure.

4.2.2.2 Procedure

The study followed a within-subject design. Prior to the experimental session, participants completed a series of online questionnaires to collect basic demographic information and measure dispositional empathy and autistic traits. The scales Perspective-taking and Empathic Concern of the Interpersonal Reactivity Scale (IRI; Davis, 1980) were used to assess cognitive and affective empathy, respectively. In the German version of the IRI (Paulus, 2009), each scale consists of 4 items, scored on a 5-point Likert scale ranging from 1 (never) to 5 (always). The short German version of the Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001; Freitag et al., 2007) was used to assess individual differences in autistic

traits. This version consists of 33 statements scored on a 4-point scale from "definitely agree" to "definitely disagree". All questionnaires were implemented in SoSci Survey (Leiner, 2018).

In the experimental session, participants completed three blocks of the emotional egocentricity paradigm. The three blocks corresponded to the happy, neutral and sad conditions, and each consisted of an emotion induction procedure, followed by the emotion perception task. Before and after each emotion induction, participants rated their current mood on a 9-point scale (-4 = sad, 4 = happy). The order of happy and sad blocks was counterbalanced across participants, but neutral was always kept in the middle. Participants had a break of 5 minutes between blocks. Before the start of the first block, participants were introduced to the paradigm. To keep the experiment's aim implicit, we presented a cover story, whereby each of the subsequent blocks contained two separate experiments, the first concerned with testing a new emotion induction procedure for future research, and the second with how we perceive emotions in a face. Participants also completed a shorter version of the emotion perception task to get familiar with the paradigm. The identity of the face morphs used in the practice was different from the face identity used in the main emotion perception task.

4.2.2.3 Statistical analysis

The effectiveness of the emotion induction was checked with a 2 x 3 analysis of variance (ANOVA) on the participants' reported mood as the dependent variable, and time (pre-induction, post-induction) and emotion condition (happy, neutral, sad) as within-subject factors. The Greenhouse-Geisser correction was used where applicable, and post-hoc t-tests with Bonferroni correction were performed to characterize the significant effects. In addition to the group analysis, mood ratings were individually screened to identify participants for which emotion induction was not successful. A total of 10 participants did not show the expected pattern of mood ratings, defined as: positive mood ratings after happy induction, negative mood ratings after sad induction, and mood ratings after neutral induction lying between those in the happy and sad conditions.

Two complementary approaches were used to examine the influence of the affective state on emotion perception. First, a repeated-measures ANOVA with PSE as dependent variable and emotion condition (happy, neutral, sad) as a within-subject factor was conducted. This approach reproduced the statistical analyses typically conducted in studies that use factorial designs to test the effects of induced mood on emotion perception (e.g., Lee et al., 2008; Niedenthal et al., 2000; Schmid & Schmid Mast, 2010). Given that this analysis takes the emotion condition as a proxy of affective state, the data from the 10 participants whose mood ratings did not indicate successful emotion induction were excluded. This resulted in an analysis sample of 40 participants.

Arguably, the self-reported mood ratings are a more accurate indication of the participant's affective experience than the condition in which each PSE was measured. Therefore,

stronger evidence of the influence of the emotional state on emotion perception would be shown if post-induction mood ratings significantly predict PSEs regardless of emotion condition. To test this, we performed a linear mixed-effects analysis of the relationship between mood ratings and PSEs. As fixed effects, we included the post-induction mood ratings as our main predictor of interest, and the pre-induction mood ratings as a control covariate. As random effects, we had intercepts for participants, as well as by-participant random slopes for the effect of post-induction mood ratings. P-values for the LMMs were computed based on Satterthwaite approximation for denominator degrees of freedom.

For this second analysis approach, we did not exclude any participant based on the reported mood in each condition. However, one influential case was detected based on examination of *DFBETAS*, a standardized measure that indicates the level of influence single observations have on coefficient estimates (Nieuwenhuis et al., 2012). To obtain unbiased regression estimates, data from this participant was removed from the analysis sample and the linear mixed model was fitted a second time. Results reported below correspond to the model tested without this influential case (n = 49). Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality.

The beta coefficients estimated for the random slope in the linear mixed model reflect the extent to which the self-reported mood influenced emotion perception for each participant. We used these coefficients as individual indices of emotional egocentricity, with more negative scores indicating stronger egocentric bias. To examine whether individual differences in empathy and autistic traits are related to the degree of self-projection during emotion perception, Pearson correlations were conducted between emotional egocentricity scores and the IRI and AQ scores.

Data and code to reproduce the statistical analyses are available at https://osf.io/5f4vn/. All statistical analyses were run in R (R Core Team, 2020) and R studio (RStudio Team, 2019). We used the following R packages: *ez* for ANOVA (Lawrence, 2016); *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017) and *influence.ME* (Nieuwenhuis et al., 2012) for linear mixed-effects analysis; *Hmisc* (Harrell & Dupont, 2018) for correlations; and *ggplot2* (Wickham, 2009) for figures.

4.3 Results

4.3.1 Manipulation check

The repeated-measures ANOVA on mood ratings revealed a significant interaction between emotion condition and time, $F(2, 98) = 126.32$, $p < .001$, $\eta_p^2 = .721$. Planned pairwise comparisons confirmed that, while there were no significant differences in mood ratings before emotion induction (all $p > .16$; pre-happy induction: $M = 1.44$, $SE = 0.09$; pre-neutral induc-

tion: M = 1.14, SE = 0.13; pre-sad induction: M = 1.12, SE = 0.12), participants felt significantly happier after happy induction (M = 2.72, SE = 0.13) than after neutral induction (M = 0.60, SE = 0.11), $t(49)$ = 12.43, p < .001, 95% CI [1.78, 2.46], Cohen's d = 1.76, and significantly less happy after sad induction (M = -1.60, SE = 0.16) compared to neutral induction, $t(49)$ = -12.66, p < .001, 95% CI [-2.55, -1.85], Cohen's d = 1.79, as well compared to happy induction, $t(49)$ = 18.41, p < .001, 95% CI [3.85, 4.79], Cohen's d = 2.60 (Figure 4.2).

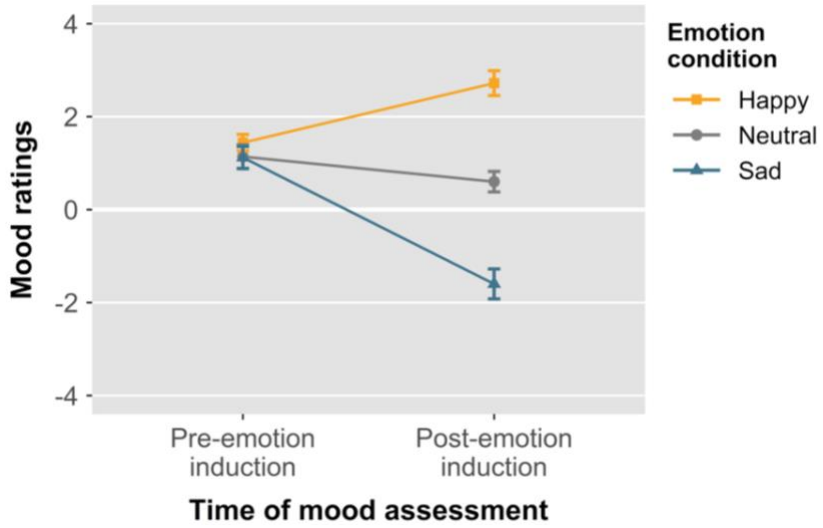

Figure 4.2. Mean self-reported mood ratings before and after each emotion induction procedure. Positive ratings represent happy state; negative ratings indicate sad state. Error bars represent within-subject 95% confidence intervals.

4.3.2 Egocentric biases in emotion perception

The ANOVA on the estimated PSEs showed a significant main effect of emotion condition, $F(2, 78)$ = 4.67, p = .018, η_p^2 = .107, indicating that emotion perception was influenced by the participants' affective state (Figure 4.3). In line with the predicted mood-congruency bias, the PSE in the happy condition (M = 67.38, SE = 0.78) was descriptively lower than in the sad condition (M = 70.45, SE = 0.82), although this difference did not reach statistical significance after correction for multiple comparisons, $t(39)$ = -2.50, p = .051, 95% CI [-5.57, -0.58], Cohen's d = 0.39. Differences in PSE between the neutral (M = 68.47, SE = 0.53) and sad conditions, and between the neutral and happy conditions were also not statistically significant (all p > .08). Though weaker, the main effect of emotion condition on PSEs remained significant even with inclusion of the full sample, $F(2, 98)$ = 3.31, p = .049, η_p^2 = .063.

The influence of the affective state on emotion perception was further demonstrated by the linear mixed model analysis, which showed that the mood participants reported after the emotion induction was a significant predictor of their PSE (Estimate = -0.40, SE = 0.18, t = -2.23, p = .031, 95% CI [-0.76, -0.05]). As indicated by the negative slope, the happier participants were, the more likely they were to judge the emotional expressions as happy, indicated

by lower PSEs (Figure 4.3). Pre-induction mood ratings did not significantly influence the PSEs (Estimate = -0.30, SE = 0.42, t = -0.71, p = .48, 95% *CI* [-1.12, -0.53]).

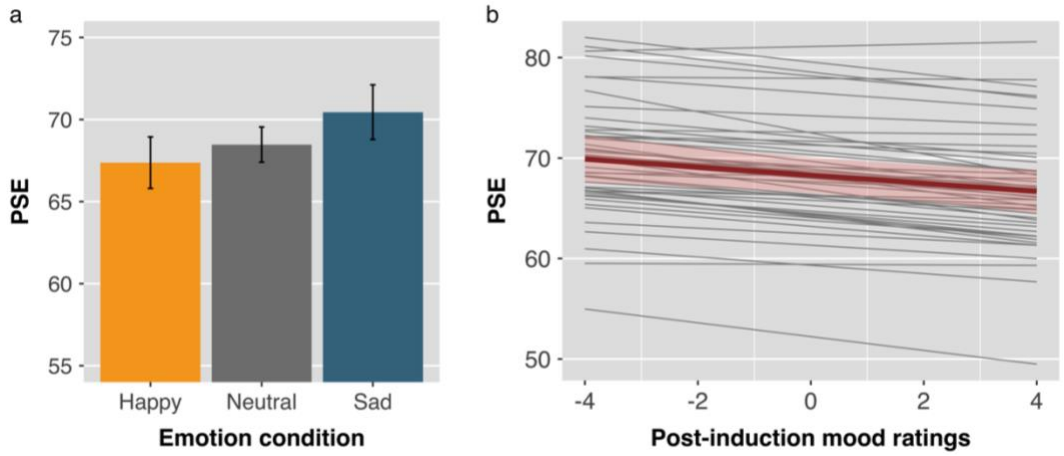

Figure 4.3. (a) Main effect of emotion condition on the point of subjective equality (PSE). PSEs indicate the percentage of happiness in the morph level at which participants were equally likely to judge the face as happy or sad. Lower PSEs are interpreted as a higher tendency to perceive happy expressions. Error bars represent within-subject 95% confidence intervals. (b) Effect of post-emotion induction mood on PSEs as estimated by the linear mixed model. The thick red regression line represents the predicted overall effect of mood (fixed effect), with 95% confidence intervals. A negative slope indicates that a happier mood predicted a higher tendency to perceive the faces as happy (i.e. lower PSEs). The estimated regressions lines for each participant (random effects) are represented with thinner grey lines.

4.3.3 Gender differences

Given that both male and female participants made emotion judgments of a single female identity, we explored gender differences by re-running the main ANOVA with gender (2 levels: female, male) added as a between-subject factor. As before, there was a significant main effect of emotion condition, $F(2, 76) = 4.85$, p = .016, η_p^2 = .113, but neither the main effect of gender (p = .27), nor its interaction with emotion condition (p = .50), were statistically significant.

Adding gender and its interaction with post-induction mood ratings as predictors in linear mixed model analysis led to the same pattern of results: post-induction mood ratings remained a significant predictor of PSE (Estimate = -0.61, SE = 0.25, t = -2.44, p = .019, 95% *CI* [-1.09, -0.12]), while gender (p = .13), its interaction with post-induction mood (p = .65), and the main effect of pre-induction mood ratings (p = .94) were not statistically significant.

4.3.4 Associations between egocentric bias, empathy and autistic traits

The slopes estimated for each participant in the linear mixed model were used as an index of emotional egocentricity, with more negative values indicating stronger egocentric bias. A weak but significant correlation was found between perspective-taking and emotional ego-

centricity bias, $r(49) = 0.36$, $p = .012$, indicating that higher perspective-taking abilities are associated with a reduced influence of the affective state on emotion perception (Figure 4.4). Neither empathic concern nor autistic traits correlated significantly with emotional egocentricity scores (Table 4.1).

Table 4.1. Summary of correlations, means and standard deviations for emotional egocentricity, IRI and AQ scores.

Variable	*M*	*SD*	1	2	3
1. Emotional egocentric bias	– 0.40	0.22			
2. Perspective-taking (IRI)	14.45	2.61	0.36* [0.08, 0.58]		
3. Empathic concern (IRI)	14.18	2.55	0.10 [– 0.19, 0.37]	0.39** [0.12, 0.61]	
4. Autistic traits (AQ)	8.61	4.80	0.01 [– 0.27, 0.29]	– 0.33* [– 0.56, – 0.05]	– 0.20 [– 0.46, 0.08]

Note. Values in square brackets indicate the 95% confidence interval for each correlation. *IRI*: Interpersonal Reactivity Index; *AQ*: Autism Spectrum Quotient; * $p < .05$; ** $p < .01$.

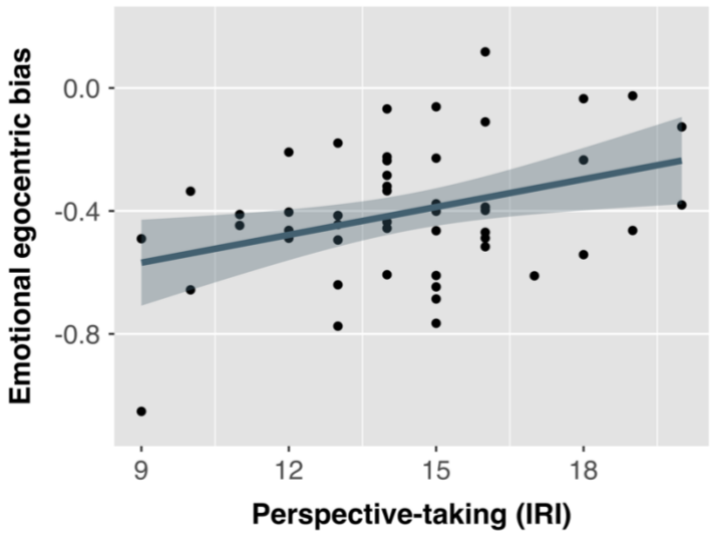

Figure 4.4. Correlation between emotional egocentricity and perspective-taking, as measured by the Interpersonal Reactivity Scale (IRI). Emotional egocentricity scores reflect the degree to which one's own emotional state biased emotion perception, with more negative scores indicating stronger egocentricity. A positive correlation indicates that emotional egocentricity decreases with higher perspective-taking tendencies.

4.4 Discussion

Research in social cognition has identified that one's own emotional experiences are an important source of information to understand how another person is feeling. Previous studies detected egocentric biases when people make inferences about someone's affective state based on information about the other's context (Hoffmann et al., 2015; Silani et al., 2013; Steinbeis & Singer, 2014). The goal of the current study was to investigate whether emotional egocentricity also occurs when affective inferences rely on reading the person's emotional expressions.

Using a combination of brief emotion induction blocks with psychophysical measures of emotion perception, we were able to detect the occurrence of egocentric biases when participants judged ambiguous emotional faces. As hypothesized, facial expressions were more readily classified as happy when participants reported feeling happy compared to sad. These results indicate a tendency to project one's own affective states when making inferences about others' emotions. Moreover, we found an association between perspective-taking and the extent to which one's own mood influenced the emotion judgments, which provides evidence that these egocentric biases are related to social cognitive abilities.

Our results replicate the mood-congruent biases in emotion perception documented in the literature. One limitation of previous studies is that they could not disentangle whether the observed effects were due to the influence of the affective state experienced by the participant, or due to more general framing or priming effects, whereby exposure to affective stimuli may increase the readiness to process cues of the same valence. Evidence for perceptual biases provided in previous emotion induction studies consisted of comparing emotion recognition between groups exposed to induction of different affective states (Lawrie et al. 2019; Lee et al., 2008; Niedenthal et al., 2000, 2001; Schmid & Schmid Mast, 2010). Although most studies included manipulation checks to demonstrate that groups differed in the experienced mood, statistical analyses could not rule out the possibility that the perceptual biases were caused by the mere exposure to an affective context, beyond whether or not this elicited an affective state to the participant. For example, Aguado et al. (2018) showed that perception of emotional expressions is enhanced when faces are introduced by a statement describing events that elicit affectively congruent emotions (e.g., angry faces are recognized faster when primed with the sentence "He notices someone has vandalized his car."). Arguably, this manipulation did not lead to significant changes in the participant's affective state, yet it elicited similar perceptual biases to those in mood induction studies.

Our study addressed this limitation by additionally using a statistical model that tested whether the participant's reported mood, instead of the emotion condition, predicted the PSEs. This statistical analysis allowed us to account for the inter-subject variability in mood ratings within conditions, which increased the power to detect any influences of affective

state. The observed relationship between affective state and emotion perception (i.e. the happier the participant was, the more likely they perceived the faces as happy) provides more direct support to the existence of mood-congruency effects, and strengthens the results from the ANOVA analyses, which despite showing an overall effect of emotion condition on PSEs, post-hoc pairwise comparisons between conditions did not reach statistical significance.

The observation of egocentric biases extends previous literature on emotional egocentricity by showing that self-projection also occurs during perception-based emotion attribution. Relying on self-knowledge for understanding other's mental states can be an efficient heuristic, especially when only limited information about the other is available (Ready et al., 2000). Given that, in our study, participants had to make quick judgments about ambiguous emotional expressions, one's own affect may have been used to guide their decisions. In fact, judgments of ambiguous faces, rather than clearer emotional expressions, have been shown to be more influenced by the participant's affective state (Cavanagh & Geisler, 2006). Future studies should assess to what extent individuals attribute their own affective states when making emotion judgments in more naturalistic situations in which additional contextual information and more time to correct for egocentric projections are available.

Importantly, the processes underlying egocentric biases in this study may be distinct from those previously observed in classic emotional egocentricity paradigms. In those tasks, participants are asked to make emotion judgments about a person based on information about the affective stimulation the other is exposed to, while at the same time receiving an affectively (in-)congruent stimulation themselves (e.g., Silani et al., 2013; von Mohr et al., 2019). To unbiasedly infer the other's feelings, participants need to disengage from their own experience and adopt the other's perspective, a process that relies on self-other distinction abilities (Lamm et al., 2016; Steinbeis, 2016). Under those conditions, emotional egocentric biases have been interpreted as a failure to differentiate between the representations of one's own affective states and others (Hoffmann, Koehne, et al., 2016; Silani et al., 2013; Tomova et al., 2014).

In our paradigm, participants were not primed to think about their own affective state while judging the other's emotional expressions, nor were asked to switch between self and other processing. In this context, the intrusion of the self-affect was more implicit and unconscious. Participants may have not actively tried to inhibit their own state nor tease apart the self- and other-representations as in previous emotional egocentricity tasks. Therefore, while our results support the idea that one's own experience is recruited when reading others' emotions, egocentric biases here should not be interpreted as an index of the participants' self-other distinction abilities.

Instead, egocentric judgments during emotion reading could be related to the participants' general disposition to shift attention towards the other's experience during social interactions. Specifically, we found a small but significant correlation between emotional

egocentricity and individual differences in the perspective-taking scale of the IRI, such that individuals with a higher predisposition to adopt the point of view of others were less influenced by their own affective states. This goes in line with the results of a recent meta-analytic study that found a positive association between dispositional perspective-taking and emotion recognition accuracy (Israelashvili et al., 2019). A higher tendency to engage in perspective-taking may lead to more attention deployed to the other during processes of mental state inference, thus minimizing the influence of one's own affect and facilitating a more accurate representation of the other's experience. Supporting this hypothesis, perspective-taking tendencies have been related to the extent to which participants focus on the other person's perspective relative to their own in a visual perspective-taking task (Bukowski & Samson, 2017). Egocentric biases in our study may have partly resulted from a lower disposition to amplify the other's representation rather than a failure in inhibiting the self-representation.

While our results suggest that components of cognitive empathy are linked to the tendency to project one's own emotions onto others, we did not find evidence for an involvement of affective empathy. Specifically, individual differences in empathic concern did not significantly correlate with the degree of emotional egocentricity. Though conclusions from null results should be drawn cautiously, our finding parallels previous studies in which emotional egocentricity was also not associated with empathic concern (Hoffmann, Banzhaf, et al., 2016) nor with other factors related to affect sharing such as alexithymia (von Mohr et al., 2019). Instead, emotional empathy seems to be more related to altercentric biases, that is, to the influence of the other's emotions on the judgment of our own affective states (Hoffmann, Banzhaf, et al., 2016).

Finally, we did not find evidence for an association between emotional egocentricity and autistic traits. Even though stronger egocentric biases have been reported in ASC (Bradford et al., 2018; Pearson et al., 2013), as well as in individuals from the general population with high autistic traits (Brunyé et al., 2012), these were mainly detected in cognitive mentalizing tasks. The only study known to us that specifically investigated emotional egocentricity did not find differences in the magnitude of egocentric biases between individuals with ASC and controls (Hoffmann, Koehne, et al., 2016). Taken together, these findings point to a dissociation in the impact of autistic traits on cognitive and affective mentalizing: while people with high autistic traits may show difficulties when inferring others' knowledge and beliefs, autistic traits do not seem to significantly impact the capacity to overcome self-projections during emotion inferences.

Some limitations of the study should be mentioned. First, only face morphs of one female identity were used in the emotion perception task. Previous research has shown that factors such as liking or the perceived similarity with the other influence the degree to which people project their own mental states (Davis, 2017; O'Brien & Ellsworth, 2012). As such, one could

expect that female participants in our study would have shown stronger egocentric biases than male participants as they were rating same-gender targets. Our exploratory analysis did not show significant gender differences in the influence of the affective state on emotion judgments. Nevertheless, given our sample size, this result should be interpreted with caution, as we had limited power to test the moderation by gender. Moreover, due to our task design, we could not have distinguished whether potential gender effects would reflect in-group biases, different motivational attitudes towards the target, or general gender differences in emotion attribution. By including more diversity of face identities and controlling for the perceived similarity with the target, future studies should address this issue and improve the generalizability of the findings.

Second, participants made binary emotion choices in the emotion perception task, while some of the morphs may have actually been perceived neither as happy nor sad. The 2-alternative forced-choice design was chosen to increase the likelihood to detect egocentricity effects and limit the task duration. To increase ecological validity, new emotional egocentricity paradigms could offer a wider range of emotion options. In addition, future task designs would benefit from adding a control condition in which participants make non-emotion judgments about the target while being under the emotion induction effects. This would allow drawing more definite conclusions about the specificity of the mood effects in relation to emotion reading vs other mood-congruent perceptual biases unrelated to social cognitive processes.

Finally, although the emotion induction procedure worked on average, there was variability in the effectiveness of the manipulation, with some participants not reporting the expected mood in each condition. Moreover, even though the assessment of PSEs was completed within a short time (approx. 1.5 min) right after each emotion induction, we cannot rule out that the strength of the induced affect decreased throughout the task. Attenuated affective self-representations may have reduced the chance of observing a bias. In future work, the use of personalized and longer-lasting forms of emotion induction could facilitate the detection of mood-congruency effects.

Notwithstanding these limitations, our study implemented a novel approach to quantify egocentric biases during emotion attribution. The adaptive psychophysical task allowed us to detect subtle changes in the tendency to perceive emotional faces as happy when participants were in different affective states. Unlike some of the previous mood-congruence studies (e.g., Harris et al., 2016; Lee et al., 2008; Niedenthal et al., 2000, 2001; Schmid & Schmid Mast, 2010), we used a within-subject design, which gave us the possibility to estimate individual bias scores and explore associations with socioemotional traits. The simplicity of the paradigm makes it appropriate to be used in emotional egocentricity research with clinical samples.

In conclusion, the current study established the existence of egocentric biases when reading facial expressions of emotion. We showed that individuals are more likely to perceive ambiguous happy-sad face morphs as happy when they are feeling happy compared to when they are sad. More importantly, the magnitude of the egocentric bias was associated with perspective-taking tendencies, which suggests that socio-cognitive processes may underlie mood-congruency biases in emotion perception. Our study extends the literature on emotional egocentricity by showing that self-projection also occurs when we rely on the other's emotional expressions for understanding their affective state.

5. Study 4: Emotional egocentricity & similarity

BRIEF REPORT

Projecting one's own affective states onto others: no influence of perceived similarity

Irene Trilla, Friedrich Eiserbeck and Isabel Dziobek

Abstract. Research has shown that people project their own emotional experiences when inferring the affective states of others. This preregistered study investigated whether the degree of affective self-projection depends on the perceived similarity with the other, an effect that has been observed for inferences of cognitive mental states. In a within-subject experiment, the perceived similarity between the participants and a series of targets was manipulated using a minimal group paradigm. After being induced positive and negative affective states, participants made emotion judgments of morphed happy and sad facial expressions of a similar and a dissimilar target. Confirming the occurrence of affective self-projection, facial expressions were judged as happier after positive induction than after negative induction. However, contrarily to more cognitive forms of social projection, we did not find evidence to support that the degree of self-projection depends on the perceived similarity with the other during quick emotion judgments.

This manuscript is available as a preprint and has been submitted for publication:

Trilla, I., Eiserbeck, F., & Dziobek, I. (2020). *Projecting one's own affective states onto others: No influence of perceived similarity*. PsyArXiv. https://doi.org/10.31234/osf.io/j2ct4

5.1 Introduction

One common strategy to understand others' mental states is to use one's own thoughts and feelings as a reference, a process referred to as self-projection. Evidence of self-projection comes from studies in which people's beliefs and affective states interfered with their inferences about the experiences of another person (Mitchell, 2009; Samson et al., 2010; Silani et al., 2013; Trilla, Weigand, et al., 2020). This is typically shown by egocentrically-biased judgments about the other's mental and affective states.

Research in social psychology has shown that individuals preferentially rely on their own experiences when making social inferences about people they perceive similar to themselves, such as in-group members (e.g., Ames, 2004b; Clement & Krueger, 2002; Davis, 2017; Robbins & Krueger, 2005). To predict the mental states of dissimilar others, people tend to recruit other inferential strategies, such as using implicit beliefs or stereotypes about a particular social group (Ames, 2004b; Ames et al., 2012), or taking as reference a known person who is more similar to the target (Willard & Markman, 2017). This flexible use of self-projection according to the perceived similarity with the other may guarantee more accurate inferences. Provided that people who share specific characteristics tend to react more alike, attributing their thoughts and feelings onto similar individuals can be an effective inferential strategy, especially when we have little information about the other. Conversely, the use of self-projection to infer the mental states of dissimilar others, or of those who are in a situation incongruent with our own current experience, can lead to stronger egocentric biases (Samson et al., 2010; Silani et al., 2013; Tamir & Mitchell, 2013). In this case, a process of adjustment is needed to correct away from the self-projection and account for the dissimilarities between oneself and the other (Tamir & Mitchell, 2013).

So far, the role of perceived similarity has been mainly demonstrated for the projection of cognitive mental states, such as traits, beliefs and attitudes (Ames, 2004b; Ames et al., 2012; Clement & Krueger, 2002; Davis, 2017). To our knowledge, only one study has assessed how similarity influences affective self-projection (O'Brien & Ellsworth, 2012). In this study, participants were deprived of drinking water, or exposed to a cold setting, before inferring the visceral states of a protagonist of a written story. Participants attributed their feelings of coldness and thirst more strongly to protagonists with whom they shared the same political views, compared to protagonists who held opposing attitudes. The study by O'Brien and Ellsworth (2012) provides initial evidence that perceived similarity influences processes of affective self-projection. However, it is yet to be determined whether such modulation extends to more automatic and low-level forms of emotion inferences. As in the case of cognitive inferences, research has shown that people tend to over-attribute their affective states when reading others' overt emotional expressions (Trilla, Weigand, et al., 2020).

To expand our understanding of the role of perceived similarity on affective self-projection, the present study examined the extent to which individuals attribute their affective states when judging emotional facial expressions of similar and dissimilar targets. Similarity was manipulated using a minimal group paradigm that assigned the same or a different arbitrary characteristic to the participant and the evaluated targets. We predicted that emotion attributions would be more biased towards one's own affective experiences when judging a similar target compared to a dissimilar target.

5.2 Methods

5.2.1 Participants

Sixty-four adults (31 females, 32 males, 1 non-binary; Age: M = 27, SD = 5.2) were recruited for this study. The sample size was decided based on a power analysis to detect a small-to-moderate effect size (Cohen's d = 0.35) using a one-tailed paired-samples t-test for the main contrast of interest, with alpha = 0.05 and power = 0.80 (see preregistration: https://osf.io/kjb84). None of the participants reported current psychiatric or neurological disorders, current psychoactive medication, severe cognitive or neuropsychological impairments, history of regular substance use, or colour-blindness.

The study conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki) and was approved by the Ethics Committee of the Psychology Department at Humboldt-Universität zu Berlin. All participants provided informed written consent.

5.2.2 Materials and procedure

5.2.2.1 Similarity manipulation

A minimal group paradigm based on the one used by Montalan et al. (2012) was conducted to manipulate the similarity between the participant and four target faces. First, participants were asked to estimate the number of dots in 10 stimulus patterns, each presented for 1000 ms. Their estimation performance was ostensibly used to determine the participant's cognitive style (*Overestimator* or *Underestimator*). In reality, the assignment of the cognitive style was pre-set and counterbalanced across participants. Participants were informed about their alleged estimation performance and read a brief description of the two cognitive styles. Overestimators were described as individuals who tend to process numerical information with a stronger focus on quantitative characteristics. Underestimators were described as individuals who tend to process numerical information focusing on qualitative characteristics.

Next, participants were asked to learn the cognitive styles of four different target faces. Two of the targets were assigned the same cognitive style as the participant ("similar targets"), and two had the other cognitive style ("dissimilar targets"). The assignment of the similarity

condition to the targets was randomized for each participant. To reduce the influence of non-manipulated factors on the perceived (dis)similarity with the targets, images of four men were used for male participants, and images of four women were used for female participants. The gender of the targets was chosen randomly for participants identified as non-binary. One similar and one dissimilar target would be later presented in the affective self-projection paradigm, and the remaining two targets, in a cognitive self-projection control task.

Participants learnt the targets' cognitive styles through an associative learning task. On each trial, a neutral face of one of the four targets, or the word "You", was presented for 1000 ms. Face stimuli were taken from the FACES database (Ebner 2010; see supplementary material for more details about the stimuli). Next, the response options ("Underestimator", "Overestimator") appeared below the stimulus, on the left and right side of the screen. Participants had to guess the target's cognitive style by pressing a left or a right keyboard key. In "You" trials, participants had to indicate their own cognitive style. The selected response was highlighted for 500 ms, after which feedback ("Correct" or "False, this person is an [over /under]estimator") appeared below the stimulus for 1000 ms. A blank screen was displayed for 500 ms before the onset of the next trial.

Participants completed a minimum of 10 blocks of 5-trials each. In each block, the four target faces, and "You", were presented once in random order. To ensure that participants assimilated the cognitive style of all targets, they completed additional blocks until each stimulus was classified correctly on 8 out of the last 10 times presented, or if a maximum of 30 blocks was reached. On average, participants completed 12.4 blocks (*SD* = 4.6). Data from two participants who did not achieve this learning criterion by the end of the task were excluded from the statistical analyses.

As a manipulation check, participants indicated their perceived similarity with each target before and after the minimal group paradigm using a 9-point scale (1 = Not similar to me; 9 = Similar to me). Participants also rated how much they liked each target (1 = I do not like this person; 9 = I like this person), and how sympathetic they found them (1 = Not sympathetic; 9 = Sympathetic).

5.2.2.2 Affective self-projection

A paradigm similar to Trilla et al. (2020) was used to estimate egocentric biases during emotion judgments of a similar and a dissimilar target. In two separate within-subject blocks, participants were first induced positive or negative affective states using audio-visual stimuli. Specifically, participants watched 1-minute videos that displayed a joyful (e.g., playful panda bears) or a sad (e.g., a dog seemingly mourning the death of another dog) animal scene, with instrumental music of matching valence playing in the background. Further details about the affect induction stimuli and procedure are available in the supplementary

material. The order of the affect induction (positive affect, negative affect) was counterbalanced across participants. To assess the effectiveness of the affective state manipulations, participants reported their current mood on an 11-point scale (1 = negative mood, indicated by a sad emoticon; 11 = positive mood, indicated by a happy emoticon) after watching the videos, and at the end of each block.

Following each affect induction, participants were asked to make quick emotion judgments of facial expressions of a similar and a dissimilar target. Ambiguous emotional expressions were created for each target by mixing an image of a happy face with a sad face at varying degrees (see supplementary material). Each trial of the emotion judgement task began with a fixation cross displayed for 1500 ms. Next, one of the morphed emotional faces of the similar or the dissimilar target was presented for 500 ms. At the stimulus offset, the response options ("happy", "sad") appeared on the right and left side of the screen (location randomly determined for each participant) until the participant made a response via a keypress.

The morph level presented in each trial was selected based on the participant's previous response following a 1-up/1-down procedure (see supplementary material). This adaptive psychophysical method allowed to reduce the number of trials needed to estimate the participant's point of subjective equality (PSE), which represents the point on a happy-sad morph continuum at which a facial expression is equally likely judged as happy or as sad. Individual PSEs were estimated for similar and dissimilar targets after each affect induction and were used as the primary dependent variable. According to the morph coding used in this study, lower PSE values indicated a higher tendency to evaluate emotional expressions as happy (vs sad). Based on this, egocentric biases would be shown if PSEs in the positive affect condition were lower than in the negative affect condition. On average, participants completed 63.4 trials (*SD* = 9.4) per block, which corresponds to approximately 2.6 minutes.

5.2.2.3 Cognitive self-projection

As a control task, we tested self-projection in the cognitive domain using a procedure based on Tamir and Mitchell (2013). This task consisted of three blocks. In the self-block, participants were asked to indicate their attitude towards a series of 30 statements (e.g., "Stormy weather is scary."; see supplementary materials for the complete item list). In the similar- and dissimilar-block, participants had to infer how a similar and dissimilar target would respond to the same statements. All ratings were made on a 9-point agreement scale (1 = Disagree; 9 = Agree). A face with a neutral expression was displayed above each statement in the similar- and dissimilar-blocks to indicate the target of the judgment. No reference was made to their cognitive style. Each statement was presented 2000 ms before the rating scale and remained onscreen until the participant responded via mouse click on the rating scale. The order of the statements was randomized. There was no response time limit, but participants were encouraged to answer as quickly as possible without overthinking their response.

To prevent participants from remembering their own ratings, the self-block was completed at the beginning of the experimental session, and the similar- and dissimilar-blocks were conducted after the affective self-projection paradigm, with counterbalanced ordered.

5.2.2.4 Procedure

The study was introduced to the participants as an investigation on how people make social inferences based on minimal information. After signing the informed consent form, participants were seated in front of the computer where they would complete all the measurements, at approximately 65 cm distance of the monitor. All tasks were programmed and presented using MATLAB R2015b (The MathWorks, Inc., Natick, Massachusetts, United States) and the Psychophysics Toolbox extension (Brainard, 1997; Kleiner et al., 2007). The supplementary Figure S5.3 shows the sequence of the tasks conducted during the experimental session. Instructions for each task were displayed onscreen at the start of each paradigm and were also reviewed orally with the experimenter to make sure they were understood.

The session began with the first part of the cognitive self-projection paradigm (self-block), in which participants had to indicate their own agreement with a series of statements. Next, participants underwent the minimal group procedure. This started with baseline ratings on similarity, likeability and sympathy for each target face. Participants then completed the dot estimation task with the pretext of finding out their cognitive style. At the end of the task, the participants alleged estimation performance ('overestimator', 'underestimator') was displayed onscreen, together with the description of each of the two cognitive styles (see supplementary material). Following the assignment of their cognitive style, participants performed the learning task to assimilate the cognitive styles of the targets. Post-learning face ratings on similarity, likeability and sympathy were completed once the participant had achieved the learning criterion (i.e. classified correctly each target face in 8 out of the last 10 times presented), or when the maximum number of learning trials had been reached.

After the minimal group paradigm, participants were introduced to the two-part procedure to assess affective self-projection. To keep the aim of the paradigm implicit, the emotional videos and the emotion judgment task were presented as two separate experiments; the first concerned with testing a new affect induction procedure, and the second with how we infer emotions from facial expressions. Participants wore headphones to listen to the music of the emotional videos, which continued playing in the background during the emotion judgment trials. The volume of the music was adjusted individually for each participant. Prior to the affect induction blocks, participants completed a practice run of the emotion judgment task, which was used to estimate their baseline PSEs. The baseline PSE defined the starting morph level presented in the adaptive procedure of the positive and negative affect blocks (see supplementary material). The order of the blocks was counterbalanced across participants,

and each consisted of the corresponding affect manipulation, followed by the emotion judgment trials. To prevent carry-over effects of the affect induction, a 5-minutes break was introduced between the two blocks, during which participants solved paper-based mazes.

The session continued with the remaining two blocks of the cognitive self-projection task. In each block, participants were asked to predict how a similar or a dissimilar target would respond to the same statements that they had rated at the beginning of the session. The order of the similar target and dissimilar target blocks was counterbalanced.

Additional measurements for potential exploratory analyses were performed at the end of the session. These included an imitation-inhibition task similar to Catmur and Heyes (2011), as well as questionnaires to collect basic demographic information (age, gender, level of education, occupation), and to measure dispositional empathy (Basic Empathy Scale; Heynen et al., 2016; and Perspective-taking and Empathic concern scales of the Interpersonal Reactivity Scale; Paulus, 2009) and autistic traits (Autism Spectrum Quotient – short version; Freitag et al., 2007). Participants were also asked to estimate the percentage of overestimators and underestimators in the population, and to report how desirable they thought each cognitive style is (1 = not desirable, 9 = desirable). None of these measurements were included for the analyses presented in this manuscript, but data is available in the OSF repository (https://osf.io/s5gp6/).

At the end of the session, participants were disclosed about the cognitive styles used in the study and were compensated with 8€/hour or course credit for their participation. On average, the experimental session lasted around 1.5 hours.

5.3 Results

Inferences were based on null hypothesis significance testing with $p < .05$. The Holm-Bonferroni adjustment was used to correct for multiple testing in post-hoc tests following significant interaction effects. Unless prespecified, all statistical tests were two-tailed. Data and R code to reproduce all analyses reported here are available at: https://osf.io/s5gp6/.

5.3.1 Similarity manipulation

To check whether the minimal group paradigm succeeded in manipulating the participants' perceived similarity with the targets, a repeated-measures analysis of variance (ANOVA) was conducted on the similarity ratings, with targets' *Similarity* (similar targets, dissimilar targets) and *Time* (pre-learning, post-learning) as within-subject factors. The ANOVA yielded significant main effects of *Similarity*, $F(1,61) = 13.34$, $MSE = 1.70$, $p < .001$, $\eta^2_p = .179$, and *Time*, $F(1,61) = 15.75$, $MSE = 0.80$, $p < .001$, $\eta^2_p = .205$, which were qualified by a significant two-way interaction, $F(1,61) = 15.60$, $MSE = 1.53$, $p < .001$, $\eta^2_p = .204$ (supplementary Figure S5.1). Follow-up pairwise comparisons confirmed that, after learning the targets' cognitive

styles, those paired with the same cognitive style as the participant were perceived as more similar to themselves (M = 5.19, SD = 1.81) than targets with the opposite cognitive style (M = 3.96, SD = 1.38), $t(61) = 4.65$, $p < .001$, 95% CI = [0.70, 1.75], Cohen's d = 0.59. No significant differences in perceived similarity were found between the targets at baseline, $t(61) = -0.09$, $p = .93$, 95% CI = [-0.39, 0.36], Cohen's d = -0.01.

A significant *Similarity*-by-*Time* interaction was also found on a repeated-measures ANOVA on likability ratings, $F(1,61) = 16.94$, $MSE = 0.94$, $p < .001$, $\eta^2_p = .217$, and on sympathy ratings, $F(1,61) = 10.80$, $MSE = 0.97$, $p = .002$, $\eta^2_p = .150$. Follow-up analyses showed that similar targets were rated as more sympathetic than dissimilar targets after learning their cognitive styles, $t(61) = 2.46$, $p = .05$, 95% CI = [0.12, 1.18], Cohen's d = 0.31, and were also liked more, $t(61) = 3.05$, $p = .01$, 95% CI = [0.29, 1.38], Cohen's d = 0.39. No significant differences on neither sympathy nor likability ratings were found at baseline (all $p > .30$). Descriptive statistics for the face ratings are available in the supplementary Table S5.1.

5.3.2 Effects of similarity on affective self-projection

5.3.2.1 Affective state manipulation

A repeated-measures ANOVA on the mood ratings, with *Affect induction* (positive, negative) and *Time of assessment* (after affect induction, at the end of the block) as within-subject factors, revealed a significant main effect of affect induction, $F(1, 61) = 320.00$, $MSE = 5.42$, $p < .001$, $\eta^2_p = .840$. As expected, participants felt more positively throughout the positive affect block (M = 8.96, SD = 1.32) than the negative affect block (M = 3.67, SD = 1.56).

A significant interaction with *Time* indicated that the strength of the induced mood faded throughout block, $F(1,61) = 77.61$, $MSE = 0.61$, $p < .001$, $\eta^2_p = .560$ (supplementary Figure S5.2). That is, participants reported feeling more positively right after watching the happy video (M = 9.40, SD = 1.32) than at the end of the block (M = 8.52, SD = 1.46), $t(61) = 7.70$, $p < .001$, 95% CI = [0.66, 1.12], Cohen's d = 0.98; and more negatively after watching the sad video (M = 3.24, SD = 1.64) than at the end of the block (M = 4.10, SD = 1.67), $t(61) = -6.21$, $p < .001$, 95% CI = [-1.13, -0.58], Cohen's d = -0.79. Despite the decline of the induced affective state during the emotion judgments, inspection of individual mood ratings confirmed that the emotional videos elicited the expected mood patterns to all participants, so no data had to be excluded on the basis of unsuccessful affect manipulation.

5.3.2.2 Emotional egocentricity biases

To check whether the similarity with the target modulated the degree of egocentric biases in emotion judgments, we ran a repeated-measures ANOVA on the PSEs, with *Affective state* (positive, negative) and *Similarity* (similar target, dissimilar target) as within-subject factors. A significant main effect of *Affective state* confirmed the occurrence of egocentric biases,

$F(1,61) = 99.06$, $MSE = 51.97$, $p < .001$, $\eta^2_p = .619$. As depicted in Figure 5.1, participants showed a higher tendency to judge the ambiguous facial expressions as happy (indicated by lower PSEs) when they were in a positive affective state ($M = 56.73$, $SD = 11.27$) than when they were in a negative state ($M = 65.84$, $SD = 8.85$).

The predicted interaction with *Similarity*, however, was not statistically significant, $F(1, 61) = 0.45$, $MSE = 9.85$, $p = .51$, $\eta^2_p = .007$. This non-significant effect was also shown in the preregistered interaction contrast, which directly compared the bias scores (i.e. the difference in PSE between negative and positive affect conditions) for similar vs. dissimilar targets, $t(61) = 0.67$, $p = .25$ (one-tailed), 95% CI = [-0.80, ∞], Cohen's $d = 0.08$ (Table 5.1).

Table 5.1. Mean and standard deviation (SD) of the point of subjective equivalence (PSE) and bias scores estimated for each target and affect block. Positive bias scores reflect egocentric biases.

	Similar targets		Dissimilar targets	
	M	*SD*	*M*	*SD*
PSE				
Negative affect	66.14	9.93	65.55	9.49
Positive affect	56.76	12.99	56.70	11.49
Bias				
[Neg. - Pos.]	9.38	8.51	8.85	7.16

To follow-up this null result, we applied equivalence testing on the critical interaction contrast using Two One-Sided Tests (Lakens et al., 2018). This procedure allowed us to test whether egocentric biases were statistically equivalent for similar and dissimilar targets. Equivalence bounds were determined based on the preregistered smallest effect size of interest (Cohen's $d = 0.35$). A significant equivalence test, $t(61) = -2.09$, $p = 0.02$, 95% CI = [-0.80, 1.86], indicated that the egocentric biases when judging emotions of similar targets were statistically equivalent to those of dissimilar targets, thus confirming the null effect of similarity.

In addition, we calculated the Bayes factor to quantify the relative evidence for the null hypothesis (i.e. no differences in bias scores between similar and dissimilar targets) over the alternative hypothesis (i.e. stronger bias scores for similar vs dissimilar targets). Bayesian analysis was performed in JASP (JASP Team, 2020). A Bayesian paired-samples t-test with the default prior set by JASP (i.e. Cauchy distribution centred at 0 with a scale parameter of 0.707, one-sided) estimated a $BF_{01} = 3.91$. This indicates that, given the data, the null hypothesis was 3.91 times more likely than the alternative hypothesis, which is considered moderate evidence for the null (Lee & Wagenmakers, 2014).

5.3.2.3 Effects of similarity on cognitive self-projection

A linear mixed-effect model (LMM) was used to test if the participant's reported attitudes (self-ratings) predicted their inferences about the attitudes of a similar and a dissimilar target (other-ratings). The other-ratings were used as the dependent variable in the LMM. As fixed effects, we included the *Self-ratings* (continuous predictor, mean-centred), *Similarity* (categorical predictor: similar target, dissimilar target), and their interaction. The maximal random effect structure supported by the task design was reduced to achieve model convergence and avoid overparameterization. Principal component analysis was used to detect and remove the random slopes that explained zero variance. The final model included random intercepts for participant, item and target identity, as well as a by-participant random slope for self-ratings and a by-item random slope for similarity. *P*-values were calculated based on Wald-statistics approximation.

The LMM yielded a significant main effect of self-rating, $b = 0.11$, 95% CI = [0.06, 0.15], $SE = 0.02$, $t = 4.60$, $p < .001$, which was qualified by a significant interaction with similarity, $b = 0.26$, 95% CI = [0.21, 0.32], $SE = 0.03$, $t = 9.48$, $p < .001$ (Figure 5.1). A simple slopes analysis indicated that, as predicted, self-ratings significantly predicted the judgments about a similar target, $b = 0.24$, 95% CI = [0.19, 0.29], $SE = 0.03$, $t = 8.85$, $p < .001$. That is, the inferred attitudes for the similar target were positively associated with the participant's own ratings about the same statements. In contrast, no significant association between self- and other-ratings was found for dissimilar targets, $b = 0.02$, 95% CI = [-0.03, 0.07], $SE = 0.03$, $t = 0.67$, $p = .50$.

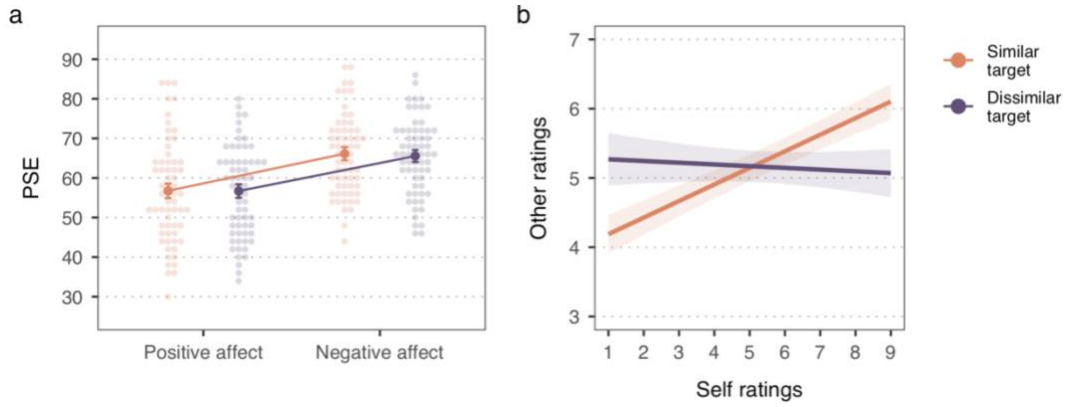

Figure 5.1. (a) Mean point of subjective equivalence (PSE) estimated for the similar and dissimilar targets in the Emotion judgment task, after positive and negative affect induction. Error bars represent within-subject 95% confidence intervals. Smaller dots depict the PSE of individual participants. (b) Association between the participant's own attitudes (self ratings) and the attitudes inferred for a similar and a dissimilar target (other ratings) in the cognitive self-projection paradigm, as estimated by the linear mixed model. Shaded areas represent 95% confidence intervals

5.4 Discussion

In line with previous studies (e.g., Niedenthal et al., 2000; Trilla, Weigand, et al., 2020), we found that judgments about others' emotional expressions are biased towards one's own affective experiences. This supports the idea that people tend to project their own affective states onto others, particularly when only limited information about the target is available (Silani et al., 2013; Trilla, Weigand, et al., 2020; von Mohr et al., 2019). However, contrary to our main prediction, the degree of emotional egocentricity did not significantly differ when judging similar targets compared to dissimilar targets. Our data was more suggestive that similarity does not modulate affective self-projection during quick emotion judgments, at least not to the extent to which it influences cognitive self-projection.

This null effect contrasts with results by O'Brien and Ellsworth (2012), who observed that participants attributed their own visceral states (e.g., thirstiness) more strongly to targets perceived as similar to themselves, compared to dissimilar others. O'Brien and Ellsworth (2012) defined similarity based on shared political attitudes. Thus, it is possible that using a somewhat arbitrary characteristic to manipulate similarity lessened the expected effects in our study. However, this is unlikely as laboratory-induced groups have been shown to elicit stronger effects on social projection than real groups (Robbins & Krueger, 2005), probably because they avoid confounding influences of pre-existing stereotypes or implicit associations. Moreover, our manipulation significantly impacted the degree of self-projection when making inferences about the targets' attitudes. Replicating previous studies (Ames, 2004b; Ames et al., 2012; Davis, 2017; Tamir & Mitchell, 2013), participants' self-ratings were predictive of the estimates made about similar targets, but not dissimilar targets. The fact that the ratings for dissimilar targets were not negatively associated with one's own attitudes suggests that self-knowledge is not used to predict the reversed attitudes for dissimilar targets (Clement & Krueger, 2002). Instead, previous studies suggest that other types of inferential strategies, such as stereotyping, are recruited when making inferences about dissimilar others (Ames, 2004b, 2004a; Ames et al., 2012; Willard & Markman, 2017).

Perhaps a more likely explanation of the null result lies in the type of emotion inferences evaluated in this study. So far, evidence of a role of perceived similarity has been found with paradigms in which participants inferred others' mental states based on information about the situation the target is experiencing (Ames, 2004a), or when predicting more stable characteristics such as attitudes and preferences (e.g., Ames, 2004b; Tamir & Mitchell, 2013). This was also the case in the study by O'Brien and Ellsworth (2012), in which participants inferred the target's visceral states based on the contextual information described in a written story. This type of social inferences requires a cognitive evaluation of the other's situation to predict the other's mental state, a process that may involve perspective-taking. In contrast, attributing affective states based on reading the other's expressive signals is a more automatic and low-level process of social inference, which could be less prone to contextual

influences. The short time frame to make the emotion judgments in our task may have also limited the chance to identify the target and recruit the learned information before their inferential process. Future studies should test whether similarity becomes a significant predictor if participants are allowed more time to process the target's identity. Moreover, the role of similarity could be further examined with emotional egocentricity paradigms in which emotion attributions result from perspective-taking processes, rather than an evaluation of the other's emotional displays (e.g., Silani et al., 2013; von Mohr et al., 2019).

Overall, results from this study strengthen the idea that individuals rely on their own beliefs and emotional experiences to infer what other people think and feel. However, the perceived similarity with the other does not seem to play a significant role in determining the degree of egocentric projections during quick emotion judgments, at least not to the same extent as it influences more evaluative and cognitive forms of social inferences.

5.5 Supplementary material

5.5.1 Supplementary figures and tables

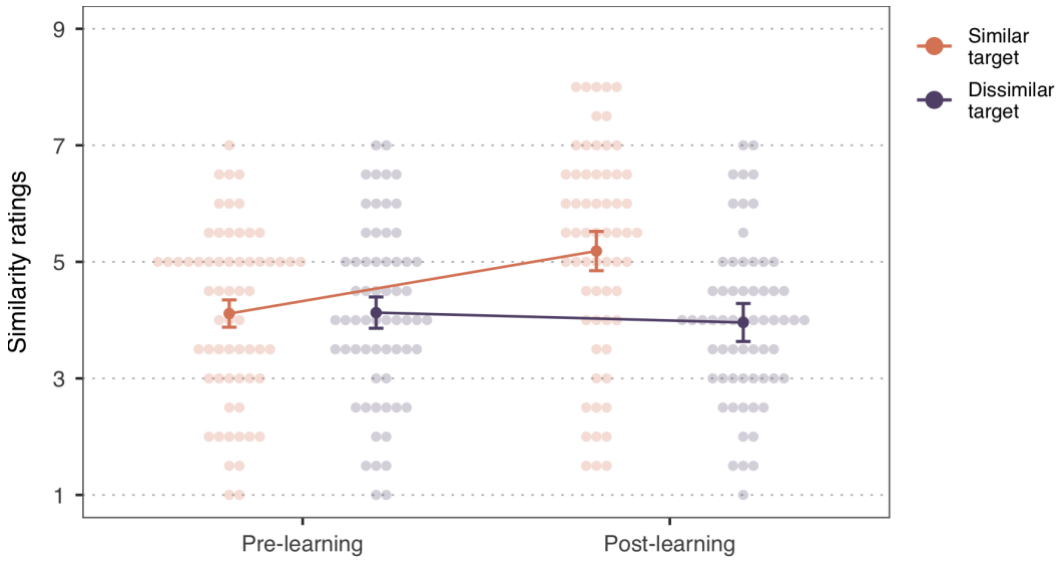

Figure S5.1. Perceived similarity with the targets, before and after learning their cognitive style. Similar targets refer to those paired with the same cognitive style as the participant; dissimilar targets are those associated with the opposite cognitive style. Error bars represent within-subject 95% confidence intervals (Morey, 2008). Smaller dots depict mean ratings of individual participants for each similarity condition and time point.

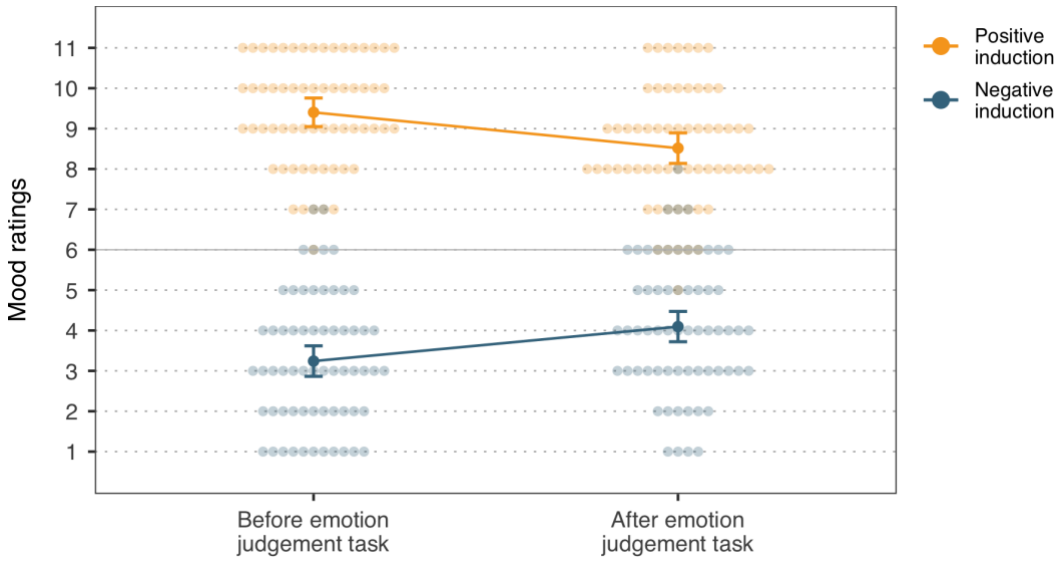

Figure S5.2. Mood ratings after each affect induction (i.e. before the emotion judgment task), and at the end of each block of the affective self-projection paradigm (i.e. after emotion judgment task). Error bars represent within-subject 95% confidence intervals (Morey, 2008). Smaller dots depict mood ratings of individual participants.

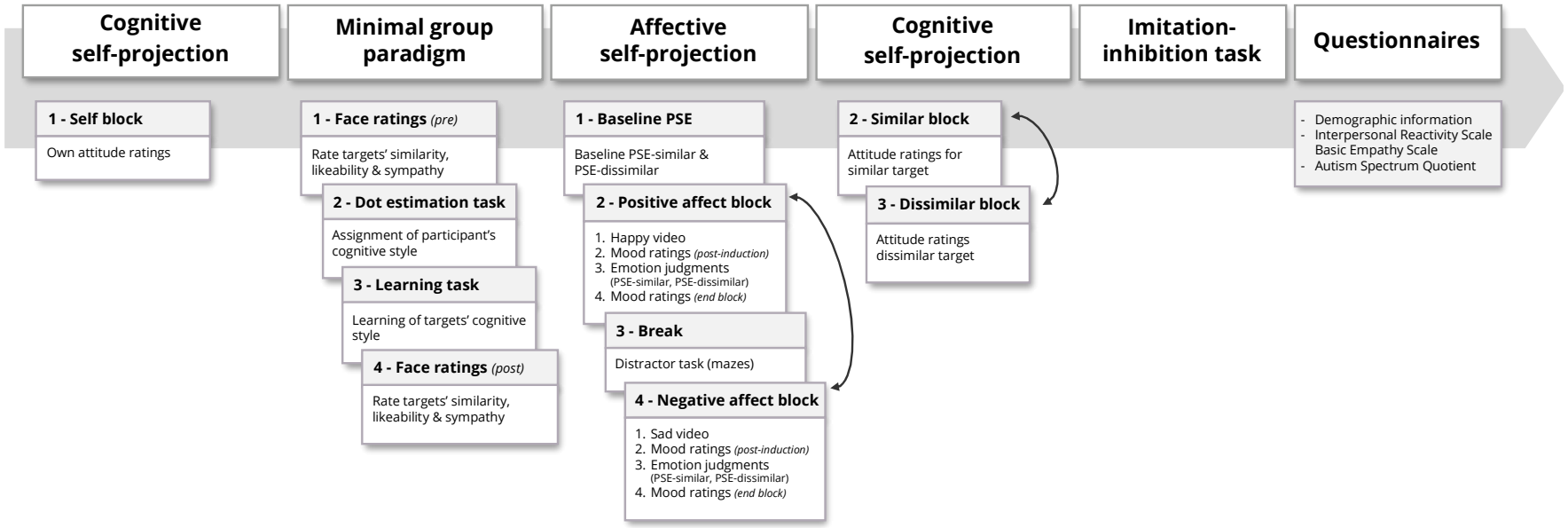

Figure S5.3. Tasks completed during the experimental session. Arrows indicate the task blocks that were counterbalanced across participants.

Table S5.1. Face ratings on similarity, likeability and sympathy. Means and standard deviations are computed for targets with the same cognitive style as the participant (similar targets) and with the opposite cognitive style (dissimilar targets), before (pre-learning) and after (post-learning) learning their cognitive style.

	Pre-learning		Post-learning	
	Similar targets	Dissimilar targets	Similar targets	Dissimilar targets
Similarity	4.11 (1.49)	4.13 (1.51)	5.19 (1.81)	3.96 (1.38)
Likeability	4.98 (1.45)	5.16 (1.44)	5.55 (1.66)	4.72 (1.40)
Sympathy	4.85 (1.33)	5.02 (1.33)	5.48 (1.57)	4.83 (1.39)

5.5.2 Minimal group paradigm: cognitive styles

Bogus cognitive styles were used to create a sense of (dis)similarity between the participant and the targets of the main experimental tasks. These cognitive styles were presumably derived from the individual's performance estimating the number of dots briefly displayed onscreen. The cognitive styles were described to the participants as follows:

German (original)	English translation
Vorangegangene Studien konnten zeigen, dass die kognitiven Profile von Personen abhängig davon, wie Sie numerische Informationen verarbeiten unterschieden werden können. Konkret können Personen als Überschätzer/in oder Unterschätzer/in klassifiziert werden.	Previous studies have shown that individuals have different cognitive profiles depending on how they process numerical information. Specifically, people can be classified as overestimators or underestimators.
Überschätzer/in: Personen mit diesem Schätzverhalten tendieren dazu, numerische Informationen mit einem stärkeren Fokus auf quantitativen Merkmalen zu verarbeiten. Im Gegensatz zu Überschätzern/innen verarbeiten Unterschätzern/innen numerische Informationen mit einem größeren Fokus auf qualitativen Merkmalen. Keines dieser Profile ist besser als das andere.	**Overestimator**: People with this estimation behavior tend to process numerical information with a stronger focus on quantitative characteristics. In contrast to overestimators, underestimators process numerical information with a greater focus on qualitative characteristics. Neither profile is better than the other.
Unterschätzer/in: Personen mit diesem Schätzverhalten tendieren dazu, numerische Informationen mit einem stärkeren Fokus auf qualitativen Merkmalen zu verarbeiten. Im Gegensatz zu Unterschätzern/innen verarbeiten Überschätzern/innen numerische Informationen mit einem größeren Fokus auf quantitativen Merkmalen. Keines dieser Profile ist besser als das andere.	**Underestimator**: People with this estimation behavior tend to process numerical information with a stronger focus on qualitative features. In contrast to underestimators, overestimators process numerical information with a greater focus on quantitative characteristics. Neither profile is better than the other.

5.5.3 Face stimuli

All face stimuli used in this study was selected from the FACES database (Ebner et al., 2010). The face models used for each task, as well as the expressions of the original stimuli, are the following:

	Minimal group paradigm	Cognitive self-projection task	Emotion judgement task
Female models	048, 054, 063, 071	063, 071	048, 054
Male models	008, 016, 057, 066	008, 057	016, 066
Facial expression	Neutral	Neutral	Happy, Sad

For the emotion judgment task, face stimuli displaying ambiguous emotional expressions were created by morphing sad and happy faces. The happy and sad faces were chosen based on normative ratings corresponding to the young subsample (Ebner et al., 2010). In particular, we used the facial expression ratings by young males to select the young male happy and sad faces, and ratings by young females to select the young female happy and sad faces. All selected emotional faces had an accuracy of at least 90% for the corresponding emotion. Using a morphing software (Morpheus Photo Morpher v3.11.), 21 morphs were generated for each model by mixing the happy face with the sad face in steps of 5%. Morphed stimuli were presented at a resolution of 235 x 302 px, which corresponds to a visual angle of 7° given a distance of 65 cm.

All face stimuli were grey-scaled and embedded within an oval to occlude the hair and clothing using GIMP (version 2.10.8, https://www.gimp.org/). To match the luminance of the images, we used two functions of the SHINE toolbox for MATLAB (Willenbockel et al., 2010): First the histMatch function was applied to the images that had the same facial identity in pairs. This function equates a set of images in terms of luminance histograms. The lumMatch function was applied to all the resulting images together to match the mean luminance and contrast.

5.5.4 Affect induction procedure and stimuli

Audio-visual stimuli were used to manipulate the participant's affective states during the affective self-projection procedure. To induce positive affect, participants watched a 1-minute video with scenes of panda bears cuddling and playing on a slide, with a cheerful instrumental song ("Get Outside!" by Jason Farnham) as soundtrack. Negative affect (sadness) was induced using a 1-minute video of a dog that was seemingly mourning the death of another dog lying on the pavement. The video was accompanied by a sad instrumental song ("Tears Won't Stop" by David Fesliyan). Participants were encouraged to allow themselves to feel the emotions the videos would elicit. We chose video scenes that portrayed

non-human animal characters instead of people to make sure that biases during emotion judgments could not be attributed to an increased sensitivity to happy or sad faces due to pre-exposure to human emotional expressions. To help maintain the induced affective state during the emotion judgment trials, the music of each video kept playing on the background until the end of the block, although the volume was automatically lowered to avoid distraction from the main task.

5.5.5 Emotion judgement task: PSE estimation

To assess biases in emotion perception, the point of subjective equality (PSE) was estimated after each affect induction (positive affect, negative affect), and for each of the two targets (similar, dissimilar). The PSE represents the point on the happy-sad morph continuum at which observers are equally likely to judge the target emotion as happy or sad. The PSE has been shown to be a sensitive index to detect biases in emotion perception (Harris et al., 2016; Marneweck et al., 2013; Trilla, Weigand, et al., 2020). In the current experiment, morph levels were coded as the percentage of happy face (vs. sad face) in the morph. For example, a PSE of 30 would indicate that a morph composed of 30% happy face + 70% sad face would be equally likely perceived as happy or sad. Consequently, lower PSEs were interpreted as a higher tendency perceive happy expressions, as it would mean that participants required less of a happy face in a morph for it to be judged as happy.

To reduce the number of trials needed to estimate the PSE, a 1-up/1-down method was implemented in MATLAB with the Palamedes Toolbox (Prins & Kingdom, 2018) following the procedure used in Trilla et al. (2020). For each target, one ascending and one descending interleaved staircases determined the morph stimulus presented on a trial-by-trial basis based on the participant's previous response. Specifically, the morph level presented in each subsequent trial within a staircase increased one step in the happy-sad morph continuum (i.e. 5% happier face) if the previous trial was a "sad" response, and decreased one step (i.e. 5% less happy face) if the previous trial was a "happy" response. Each staircase stopped after 8 reversals, which are the transition points within a staircase in which the participant switches from perceiving a happy face to a sad face, or vice versa. The PSE for each condition was estimated by averaging the morph levels of the last 5 reversals of both staircases set for that particular target and affective state condition.

To prevent order effects, blocks of trials of the similar and dissimilar target were alternated within each run of the emotion judgement task. That is, participants first completed approximately half of the trials for the similar target, then the first half of the dissimilar target, and so on. Which of the two targets was presented first was chosen randomly. To make sure that participants recognized the identity of the face morphs, an image of the target with a neutral facial expression was displayed at the beginning of each block and remained onscreen until the participant pressed a key to start.

Before the first affect induction, participants completed one block of trials for the similar target and one for the dissimilar target in order to estimate the baseline PSEs. The ascending and descending staircases in the baseline blocks started at 0% happy - 100% sad morph, and 100% happy - 0% sad morph, respectively. The PSE calculated at baseline for each target was used to define the start points of the staircases of the experimental blocks. Specifically, the ascending and descending staircases for each target started 4 morph steps lower or higher (i.e. ± 20% happier face) than the corresponding baseline PSE. The individual adjustment of the staircases' settings allowed to reduce the task duration of the main experimental blocks to assure that estimation of the critical PSEs was within the effects of the affect induction. The number of trials completed on average was 95.8 (*SD* = 14.5) at baseline, 62.6 (*SD* = 9.27) after the positive affect induction, and 64.2 (*SD* = 9.47) after the negative affect induction.

5.5.6 Cognitive self-projection paradigm: items

Items rated by participants in the cognitive self-projection paradigm, on a 9-point agreement scale (1 = Not agree; 9 = Agree).

German (original)	English translation
Alternative Heilmethoden sollten verboten werden.	Alternative healing methods should be banned.
Das Frühstück ist die wichtigste Mahlzeit des Tages.	Breakfast is the most important meal of the day.
Das Leben braucht Party.	Life needs party.
Das Öffentliche Gesundheitssystem ist gut.	Public health system is good.
Der Kapitalismus ist etwas Gutes.	Capitalism is a good thing.
Die Arbeit von Politikern ist wichtig.	The work of politicians is important.
Die meisten Leute machen sich zu viele Sorgen.	Most people worry too much.
Die Nutzung von Tieren in Zirkussen sollte erlaubt sein.	The use of animals in circuses should be allowed.
Die Vorstellung von Künstlicher Intelligenz ist gruselig.	The idea of artificial intelligence is scary.
Elektroautos sind die Zukunft.	Electric cars are the future.
Es sollte eine Geldstrafe für Nicht-Recycling geben.	Not recycling should be fined.
Essen macht müde.	Food makes you tired.
Flugzeuge sind faszinierend.	Airplanes are fascinating.
Gedichte sind wunderbar.	Poems are wonderful.
Gedruckte Bücher sind besser als eBooks.	Printed books are better than eBooks.
Hausaufgaben sollten verboten werden.	Homework should be prohibited.
Hunde sind besser als Katzen.	Dogs are better than cats.
Kuscheln ist besser als Küssen.	Cuddling is better than kissing.
Lakritz ist lecker.	Licorice is delicious.
Macs sind besser als PCs.	Macs are better than PCs.
Models sind zu dünn.	Models are too thin.
Stilles Wasser ist besser als Sprudelwasser.	Still water is better than sparkling water.
Strandurlaub ist besser als Städteurlaub.	Beach vacations are better than city vacations.
Stürmisches Wetter ist gruselig.	Stormy weather is scary.
Supermärkte sollten Sonntags offen haben.	Supermarkets should be open on Sundays.
Tanzen macht Spaß.	Dancing is fun.
Teure Dinge sind hochwertiger.	Expensive things are more valuable.
Videospiele sind zu gewalttätig.	Video games are too violent.
Weiche Drogen sollten legalisiert werden.	Soft drugs should be legalized.
Weihnachten ist ein wichtiger Feiertag.	Christmas is an important holiday.

5.5.7 R packages used for statistics

All statistical analyses were conducted in R (R Core Team, 2020) and R studio (RStudio Team, 2019), except for the calculation of Bayes factors, which was done in JASP (JASP Team, 2020). The main R packages used were: *afex* (Singmann et al., 2019) for ANOVA; *rstatix* (Kassambara, 2020) for summary statistics and t-tests; *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017) and *sjPlot* (Lüdecke, 2020) for linear mixed-effects models; *pwr* (Champely, 2018) for power calculations; *TOSTER* (Lakens, 2017) for equivalence testing; Rmisc (Hope, 2013) for calculation of within-subject confidence intervals; and *ggplot2* (Wickham, 2009) and *patchwork* (Pedersen, 2019) for figures.

The full reproducible code is available in the OSF repository: https://osf.io/s5gp6/.

6. General discussion

6.1 Part I – Altercentrism: Mimicry

6.1.1 Contextual factors modulate mimicry

Current accounts understand mimicry as a motivated and context-dependent behaviour that fosters social adaptation and affiliation (Hess & Fischer, 2013; Wang & Hamilton, 2012). Study 1 and Study 2 provide evidence favouring this notion by showing that people adjust their mimicry behaviour according to social contextual factors.

6.1.1.1 Not all emotional expressions are spontaneously mimicked

Mimicry involves integrating social and communicative cues to discern the other agent's affiliative intent. This is especially relevant in the case of facial mimicry, in which emotional expressions are an inherent part of the action to be imitated (Fischer & Hess, 2017). Expressions that signal approachability, such as smiles, have been shown to elicit facial mimicry reactions consistently (Bourgeois & Hess, 2008; Hess & Bourgeois, 2010; van der Schalk, Hawk, et al., 2011). In contrasts, expressions that signal threat or avoidance, such as an angry face, may discourage the observer from mimicking the other, as responding with a negative expression could be socially costly (Fischer & Hess, 2017). This pattern of spontaneous imitation of approach vs non-approach emotions was observed in Study 1: While happy expressions were generally mimicked, there was no evidence of congruent facial reactions when observing angry faces, at least as indicated by the activity recorded over the frowning muscle (i.e. Corrugator Supercilii).

This result adds to a growing body of studies that also failed to elicit mimicry in response to angry faces (e.g., Blocker & McIntosh, 2016; Bourgeois & Hess, 2008; Deng & Hu, 2018; Pavarini et al., 2019; Rymarczyk et al., 2011). Most of these experiments, including ours, evaluated mimicry using relatively minimal tasks that do not provide any context for the emotional expressions displayed nor define a specific interaction goal. In Study 1, participants were asked to passively view the stimuli and focus on the targets' physical attributes. Under these conditions, angry expressions may have been perceived as an antisocial signal, thus decreasing the chances for mimicry.

Future studies interested in mimicry of anger should consider experimental designs in which mimicking non-approach emotions could be advantageous for the task at hand. For example,

paradigms that ask participants to evaluate the target's emotional expressions explicitly (e.g., emotion recognition tasks) tend to elicit stronger facial mimicry of both positive and negative emotions than tasks in which participants judge non-emotional or non-social features of the target (Murata et al., 2016; van Dillen et al., 2015). Facial mimicry in contexts in which there is a motivation to decipher the other's affective state may be recruited to support emotion recognition (Drimalla et al., 2019).

Mimicry of non-approach emotions could also be triggered by setting up situations in which the expressed emotion is directed at a third party. For instance, providing contextual information that justifies an angry expression (e.g., "This person has just seen someone else hurting a child."), or presenting an angry face that directs the attention to a third party. If the observer and the target share the same anger-eliciting view, mimicry of anger may serve an affiliative goal and would, therefore, be more likely to occur (Bourgeois & Hess, 2008).

6.1.1.2 Reward modulates mimicry, but how?

In Study 1, we observed mimicry responses (i.e. increased Zygomaticus Major activity) to smiles of targets previously associated with low vs high rewards. The significant reward-driven modulation of mimicry is congruent with motivational accounts, as it shows that mimicry is dependent on contextual factors. However, the direction of the reward effect was surprising. Based on previous literature (Korb et al., 2019; Sims et al., 2012), we expected stronger facial mimicry responses to targets associated with higher rewards than targets associated with low rewards, particularly the former were evaluated more positively.

Even though this result is at odds with our predictions, this is not the first experiment that showed up-regulation of mimicry when responding to seemingly less approachable targets. For example, enhanced mimicry was reported for angry out-group members compared to in-group members (Rauchbauer et al., 2016) and agents that had previously excluded the participants (Hühnel et al., 2018; Lakin et al., 2008). These findings have been interpreted as evidence that mimicry can be flexibly regulated to support appeasement in response to counter-affiliative signals (Cheung et al., 2015; Hühnel et al., 2018; Lakin et al., 2008; Rauchbauer et al., 2016). Individuals may mimic others in negative contexts as a means to establish or regain a positive social bond.

Differences in the sample characteristics could also account for the unexpected results. In Study 1, we only included male participants and same-gender face identities. This may have changed the evaluation or type of relationship established between the observer and the targets compared to previous studies, whose samples included mostly women (Korb et al., 2019; Sims et al., 2012). Moreover, the influence of reward on mimicry has been shown to vary as a function of autistic traits, with higher autistic traits being associated with weaker or reversed reward effects (Haffey et al., 2013; Sims et al., 2012). Given that men tend to score higher than women in measures of autistic traits (Baron-Cohen et al., 2001; Freitag et al.,

2007), mimicry patterns in our study could potentially reflect the behaviours of individuals with higher autistic traits.

Nevertheless, we cannot rule out the possibility that the observed influence of reward is related to other uncontrolled factors of the experimental settings, or even a chance finding. Given the relatively low statistical power of our and prior studies, it is likely that some of the reported effects have been overestimated. In addition, it is important to disentangle the exact factors that underlie the influence of the reward manipulations on mimicry. One implied assumption in these studies is that the reward effects are due to a change in the attitude towards, or likeability, of the agents associated with high vs low probabilities of winning money (Korb et al., 2019). However, the reward manipulations could also affect other factors, such as the attentional salience (Sali et al., 2014) or the personal relevance of the stimuli (Forbes et al., 2021). Further studies are thus required to resolve the inconsistent observations regarding the influence of reward on mimicry and determine the mechanisms behind such effects.

6.1.1.3 Gaze matters, but only under certain conditions

Study 2 was planned to test further the idea that mimicry is up-regulated when interacting with a partner who signals social approachability. In this case, we assessed the role of direct and averted gaze as cues of social interest and rejection (Wirth et al., 2010). In contrast to earlier findings (Wang, Newport, et al., 2011; Wang & Hamilton, 2014), we could not detect an overall effect of gaze on mimicry. The degree of automatic imitation was not significantly different when individuals viewed targets who established direct gaze, compared to when targets averted the gaze away. This data adds to other failed replications of the gaze effects on mimicry (Farmer et al., 2021; Marsh et al., 2016). As in our case, neither of these experiments could support the original finding of stronger automatic imitation following direct eye contact vs averted gaze.

The fact that the influence of gaze had been observed in follow-up experiments (Forbes et al., 2017; Prinsen et al., 2017; Wang & Hamilton, 2014) brings confidence to the replicability of the original report. However, most of these studies were conducted by the same research group and/or used slight variations of the same paradigm and stimuli. When focusing on a theoretical prediction rather than the effects within a given paradigm, a combination of direct and conceptual replications is recommended to build confidence in a result (Brandt et al., 2014). A finding that can be replicated by other researchers in a different sample and using different experimental procedures may confirm that this is knowledge that can be separated from the specific circumstances under which it was produced. Based on the existing set of studies investigating the role of gaze on mimicry, this does not seem to be the case.

Given the complexity of the social context, random factors related to the sample, the task or the stimuli could have confounded the effects observed in the early studies. For example, the

original paradigm displayed direct and averted gaze cues by a single white female target with a smiling expression (e.g., Forbes et al., 2017; Wang, Newport, et al., 2011; Wang & Hamilton, 2014). Given that direct eye contact with a smile is often perceived as a socially rewarding cue (Niedenthal et al., 2010), the positive emotional expression in the original studies may have intensified the influence of the gaze manipulation. In addition, exploratory analyses in Study 2 suggested that gaze effects may be stronger when observing white agents compared to dark-skinned agents, at least in a predominantly white participant sample. This is congruent with literature showing an influence of group membership and social status on the perception of others' gaze cues (Collova et al., 2017; Dalmaso et al., 2012, 2015; Weisbuch et al., 2017).

Compared to prior experimental designs (e.g., Forbes et al., 2017; Wang, Newport, et al., 2011; Wang & Hamilton, 2014), Marsh et al. (2016) and Study 2 presented more diverse target identities with different ethnic backgrounds and neutral facial expression. Moreover, the statistical approach used in Study 2 (i.e. mixed-effects models) also took into account the sampling variability of the stimuli. Arguably, these experimental procedures may have counteracted some of the random factors that might have enhanced the gaze effects in earlier studies. In other words, it is possible that gaze influences mimicry, but only under specific conditions. Future work will need to systematically tease apart the impact of each contextual factor and their interactions to draw a more comprehensive picture of how and when gaze modulates mimicry.

6.1.2 Individual variability in the contextual modulation of mimicry

As discussed, mimicry is a flexible and adaptive behaviour that depends on situational factors. However, not everybody adapts to the context in the same way. Individual differences in personality traits, socio-cognitive abilities and psychobiological predispositions can determine the contextual modulation of mimicry. Studies 1 and 2 re-examined the role of two factors previously associated with differential degrees of mimicry modulation, namely oxytocin and autistic traits.

6.1.2.1 Inconclusive evidence regarding the role of oxytocin

Following initial evidence of a role of oxytocin on mimicry (De Coster et al., 2014; Korb et al., 2016; Pavarini et al., 2019), we hypothesised that intranasal oxytocin would increase facial mimicry compared to placebo, especially when observing emotional expressions of high reward vs low reward agents. Data from Study 1 were, however, inconclusive. Our statistical analyses did not yield any significant effects of oxytocin, neither with regards to a direct influence on mimicry nor on the modulatory effects of reward. To further assess the non-significant results, we conducted equivalence testing, a procedure that allows for the rejection of effects at least as large as the effect size of interest (Lakens et al., 2018). Results

suggested data insensitivity: we could neither conclude whether oxytocin had an effect, nor that the placebo and oxytocin groups behaved similarly. The study just did not have enough power to detect even a strong effect reliably.

6.1.2.1.1 Statistical power issues

Results of the equivalence tests reflected a pervasive problem within intranasal oxytocin research. Until recently, behavioural intranasal oxytocin studies typically tested a similar or lower number of participants to our experiment (e.g., De Coster et al., 2014; Korb et al., 2016; Tomova et al., 2019). Between-subject designs with 60 participants can only reliably detect large effects (Cohen's d = 0.74) with 80% power. Walum et al. (2016) estimated that approximately 80% of all attempts to detect a true effect in intranasal oxytocin studies would fail due to low statistical power. This is especially concerning for social cognition research, in which contextual factors typically exert small effects (Richard et al., 2003). Because underpowered studies also reduce the probability that a statistically significant result reflects a true effect (Button et al., 2013), it is likely that at least some of the positive findings regarding the role of oxytocin on social behaviour are false positives.

While there seems to be a trend in more recent oxytocin studies towards larger samples (e.g., Pavarini et al., 2019), estimations by Walum et al. (2016) indicate that sample sizes of hundreds of individuals are necessary to produce reliable data in social cognition research. Multi-site collaborative studies could help towards this endeavour. In addition, meta-analyses that pool the effects across experiments could help derive more precise estimates.

6.1.2.1.2 Oxytocin administration

Parallel to improving the study samples, more attention should be placed on understanding the procedures for oxytocin administration. Given the physiology of the oxytocinergic system, administrations through nasal sprays are the only available non-invasive path to manipulate oxytocin levels. However, there is still uncertainty on the optimal doses and timings for intranasal oxytocin treatments (Quintana et al., 2021). Spengler et al. (2017) identified that a time window between 45 and 70 min and a dose of 24 IU elicited the most robust neural responses to emotional stimuli. However, different administration devices, gender, timing, and factors related to the person's state of mind seem to differentially affect neural and behavioural responses (Quintana et al., 2021). It is also still unclear to what extent oxytocin exerts cognitive and behavioural effects by acting on the central nervous system, the peripheral system, or both. Without understanding the dose-response pattern and physiology of oxytocin treatments, non-significant effects may simply be due to incorrect manipulation of oxytocin levels (Quintana et al., 2021).

Taken all together, current evidence of the role of oxytocin on facial mimicry remains weak. Beyond the likelihood of false positives, oxytocin did not have a robust or consistent influ-

ence on mimicry in previous studies. The effect of intranasal oxytocin on automatic imitation originally reported (De Coster et al., 2014) was not replicated (Tomova et al., 2019); and oxytocin only influenced facial mimicry in response to certain stimuli (Korb et al., 2016), emotional expressions (Korb et al., 2016; Pavarini et al., 2019) and type of tasks (Korb et al., 2016). Nevertheless, the fact that our and prior studies could not make robust conclusions is not proof that oxytocin does not play a role in mimicry behaviour. Given its close links with socio-cognitive processes such as empathy and emotion understanding (Hurlemann et al., 2010; Shahrestani et al., 2013), and its assumed role as a regulator of behaviour (Ma et al., 2016; Quintana & Guastella, 2020), oxytocin's involvement in the neurobiology of mimicry seems still plausible. Future studies will need to overcome the methodological issues that have challenged oxytocin research to draw more reliable insights.

6.1.2.2 Individual differences in social functioning

6.1.2.2.1 Autistic traits

Previous studies had observed reduced gaze effects on mimicry in individuals with ASC (Forbes et al., 2017; Vivanti & Dissanayake, 2014). Study 2 extended these findings by showing that autistic traits are also associated with atypical social modulation of automatic imitation in a sub-clinical sample. As predicted, automatic imitation following averted gaze was stronger with increasing levels of autistic traits. However, no significant associations with autistic traits were found for mimicry responses to direct gaze. This pattern of results is congruent with previous studies on ASC. While neurotypical participants tended to mimic targets with averted gaze less than those with direct gaze, participants with ASC automatically mimicked the observed movements regardless of gaze cues (Forbes et al., 2017; Vivanti & Dissanayake, 2014).

In addition to atypical responding to gaze, individuals with high autistic traits have been shown to be less influenced by prosocial priming (Cook & Bird, 2012) and the reward value of the observed agent (Haffey et al., 2013; Sims et al., 2012) than individuals with low autistic traits. Altogether, these studies are in line with the social top-down response modulation theory, which proposes that mimicry patterns in ASC arise from a decreased adaptability of mimicry to the social context (Wang & Hamilton, 2012). Individuals with high autistic traits may imitate others' actions automatically and spontaneously, but contextual factors do not influence their behaviour to the same extent as in individuals with low autistic traits.

6.1.2.2.2 Social anxiety

According to the exploratory analyses in Study 2, other impairments in social functioning, such as social anxiety, could also be associated with particular mimicry patterns in response to direct and averted gaze cues. Specifically, higher dispositional fear of interacting with others was linked to a stronger tendency to automatically imitate the observed actions fol-

lowing averted gaze, as well as with a weaker automatic imitation following direct gaze. Descriptively, this mimicry pattern was similar to the one detected in individuals with higher autistic traits. Post-hoc analyses suggested that social anxiety accounted for a higher variance than autistic traits in predicting automatic imitation tendencies following direct and averted gaze in our sample. Given the moderately high covariation between autistic traits and social anxiety (Cath et al., 2008; Spain et al., 2018), future research will need to disentangle the relative contribution of each of the two factors on the gaze modulation of mimicry.

If a role of social anxiety and autistic traits is confirmed, a further step will be to identify the causes behind the atypical mimicry responses. The fact that two correlated factors are associated with similar behaviour does not necessarily imply common underlying mechanisms. According to the social motivation theory, the dampened social regulation of mimicry in ASC could be related to a reduced sensitivity to social stimuli and rewards (Chevallier et al., 2012). If a person is intrinsically less motivated to engage socially, cues such as eye contact and emotional expressions may be less salient or relevant for regulating one's behaviour.

In contrast, mimicry responses to direct and averted gaze in social anxiety could be related to a different evaluation of gaze cues compared to individuals without social anxiety. For socially anxious individuals, someone looking at them may not be a sign of social interest but an indication of being scrutinised (Schulze et al., 2013). Since social anxiety is characterised by a fear of being evaluated (Mattick & Clarke, 1998), direct gaze could become a signal of social threat and trigger anxiety. Individuals with social anxiety may thus inhibit mimicry when observing someone with direct gaze to avoid endorsing the social interaction.

In sum, Study 2 identified socio-emotional traits that could account for some of the inconsistent results regarding the contextual modulation of mimicry across studies. Individual differences in social functioning, either related to autistic traits, social anxiety, or a common construct, may explain why different samples or individuals adjust their mimicry behaviour to the social context differently.

6.1.3 Up- or down-regulation of mimicry?

Some studies on the social modulation of mimicry have presented the effects as an enhancement of mimicry in situations where such behaviour would lead to a positive social outcome, rather than mimicry inhibition in unfavourable conditions. For example, Wang et al. (2011) attributed the gaze effects to an up-regulation of automatic imitation following direct gaze (vs averted gaze); and Sims et al. (2012) interpreted the influence of reward as an enhancement of mimicry when observing high reward stimuli (vs low reward). However, as long as only two conditions are being compared against each other, the effects could be either reported as "stronger imitation following direct gaze relative to averted gaze" or "weaker imitation for averted gaze relative to direct gaze". How one describes these differences is just

a matter of which of the two conditions is taken as a reference, but this is an arbitrary choice without a baseline condition.

If mimicry is an automatic and spontaneous process, inhibition should be considered a primary mechanism for its regulation (Brass et al., 2009; Heyes, 2011). Data from Study 2 are consistent with this idea. First, autistic traits were primarily predictive of mimicry reactions following averted gaze; the association between autistic traits and mimicry to direct gaze did not reach significance. Second, according to the exploratory analyses, participants mimicked the dark-skinned agents irrespective of their gaze direction. In contrast, white agents were mimicked only following direct gaze. Although not tested directly, this pattern of results suggests that the interaction between the target's ethnicity and gaze was mainly driven by a reduced inhibition of mimicry to dark-skinned targets with averted gaze compared to white targets. Based on these preliminary observations, inhibitory mechanisms seem to underlie the gaze-driven modulation of mimicry.

Nevertheless, studies that included a control condition to compare the effects of each contextual manipulation have also provided evidence of mimicry enhancements. For example, Leighton et al. (2010) found up-regulation of mimicry after prosocial priming and down-regulation after antisocial priming, compared to a no priming condition. Moreover, Likowski et al. (2008) reported stronger facial mimicry when observing positive characters compared to neutral ones, but did not find significant mimicry differences between negative and neutral characters. Similarly, Wang and Hamilton (2014) reported increased automatic imitation to direct gaze compared to a no-gaze condition, but no significant differences were detected between averted and no-gaze. Considering these studies, mimicry regulation seems to involve inhibitory mechanisms but also activation of mimicry in circumstances when it may not be automatically triggered.

Research on the neural underpinnings of mimicry control could also help shed light on how the contextual modulation of mimicry is implemented. Neuroimaging studies have identified the right temporoparietal junction (TPJ) and the medial prefrontal cortex (PFC) as two critical regions for mimicry control (Brass et al., 2005, 2009). On the one hand, the right TPJ is thought to support the ability to distinguish and switch between co-activated self and other motor representations (Brass et al., 2005, 2009; Sowden & Catmur, 2015). On the other hand, the medial PFC seems to be linked to the inhibitory control of mimicry according to one's own intentions and contextual factors (Brass et al., 2009; Korb et al., 2019; Wang, Ramsey, et al., 2011; Wang & Hamilton, 2015). Crucially, both the right TPJ and medial PFC are more engaged when the tendency to automatically mimic other's movements or facial expressions conflicts with the task demands (Korb et al., 2019; Wang, Ramsey, et al., 2011; Wang & Hamilton, 2015), which supports the idea that mimicry arises automatically and spontaneously, but can be inhibited according to the social context (Wang & Hamilton, 2012).

6.1.4 Low replicability of research findings: a matter of context?

As discussed, the effects of the main contextual modulators tested (i.e. reward and gaze) were either not in the expected direction (Study 1) or not statistically significant without accounting for additional contextual or intrapersonal factors (Study 2). The fact that these experiments did not replicate previous literature may not be as surprising, considering the well-known replicability crisis in psychology (Earp & Trafimow, 2015).

Replications are essential for theoretical development through confirmation and disconfirmation of results. A single study will not provide, on its own, sufficient conclusive evidence of an effect. It is the accumulation of evidence from different studies conducted in different labs and contexts that will increase the confidence in the explored phenomena (Schmidt, 2009). While none of the mimicry experiments presented here was planned as a direct replication of the original studies, conceptual replications and follow-up experiments also play an essential role. Conceptual replications can detect if variations in methodological approaches lead to significant changes in the phenomena studied and test the generalizability of the results beyond the original method (Schmidt, 2009; Stroebe & Strack, 2014; Trafimow, 2019). For example, if we are interested in mimicry as a general phenomenon, it is important to test if the original effects are generalisable across different forms and mimicry operationalisations. Moreover, it is necessary to rule out the possibility that the observed effects are due to the idiosyncrasy of the stimuli and context of a particular study.

Failures to replicate are prevalent in research investigating highly contextual phenomena, such as social cognition. A higher contextual sensitivity of an effect makes it more likely that any small variations in the experimental conditions, including sample characteristics or task design, impact the direction of the effects (Van Bavel et al., 2016). A replication failure does not necessarily mean that the original finding is fraudulent nor that the follow-up study was incorrect. Just as replication studies may be affected by particular contexts, original results may also simply be due to the specific settings in which they were initially tested (Brandt et al., 2014; Noah et al., 2018). As such, unsuccessful replications can also be key to unlocking the underlying psychological processes of an effect (Brandt et al., 2014).

The mimicry experiments presented here might exemplify the complexity of studying context-dependent phenomena. The fact that we could not replicate the findings of previous studies highlights the need to systematically evaluate which conditions are necessary to observe an effect of contextual cues on mimicry. As in any other field or theory, insights from cumulative replication and extension studies will allow us to draw a more reliable and confident picture of the contextual modulation of mimicry.

6.2 Part II – Egocentrism: Self-projection

6.2.1 A new framework to study affective self-projection

The use of self-representations as a reference point to understand others has been a topic of interest in social psychology for many years. However, most of the literature has focused on investigating self-projection in the cognitive domain, that is, on the over-attribution of one's own beliefs, traits and attitudes onto others (e.g. Ames, 2004b; Mitchell, 2009; Ross et al., 1977), neglecting the influence of own affective states. More recently, a shift towards research on the affective domain was initiated by Silani et al. (2013), who developed a paradigm to investigate emotional egocentricity. The original results have been replicated with variations of the same affective perspective-taking paradigm and across research labs (Hoffmann, Banzhaf, et al., 2016; Hoffmann et al., 2015; Hoffmann, Koehne, et al., 2016; Riva et al., 2016; Steinbeis, 2016; von Mohr et al., 2019), which brings confidence to the occurrence of self-projection in the affective domain.

Part II of this dissertation aimed to provide a complementary framework to study emotional egocentricity. In particular, we investigated if one's own affective states also influence how we read other's emotional expressions, a more automatic and low-level process than the type of affective inferences previously researched. Supporting the occurrence of egocentricity in emotion perception, Study 3 detected that facial expressions were judged as happier after participants were induced a happy state, compared to when they were feeling sad. The egocentric influence of one's own affective states on the evaluation of others' emotional facial expressions was further replicated in Study 4.

6.2.1.1 The paradigm: lessons learnt

Emotional egocentricity in our studies was measured using a paradigm inspired by previous emotion perception research investigating mood-congruency biases (e.g., Qiao-Tasserit et al., 2017; Schmid & Schmid Mast, 2010). It consisted of two parts. First, participants underwent an emotion induction procedure to control the affective state experienced in each experimental condition. Second, participants were asked to categorise ambiguous facial expressions (i.e. morphs created by mixing a happy and a sad face at varying degrees) as either happy or sad. Using an adaptive psychophysical procedure, we were able to estimate changes in the tendency to judge the faces as happy whilst being in a happy state, compared to when feeling sad.

Overall, our paradigm proved to be a suitable procedure to measure egocentricity during processes of emotion perception. The short duration, simplicity and within-subject nature of the paradigm make it feasible to be applied in a variety of clinical and non-clinical samples, as well as to investigate individual differences in self-projection. Nevertheless, a few points should be taken into consideration in future studies when using similar procedures.

6.2.1.1.1.1 Emotion induction

Study 3 used a combination of autobiographical recall and audio-visual stimuli to induce happy, neutral and sad states. This procedure was successful in manipulating the participants' affect. However, during the debriefing of the experiment, a few participants reported that some of the images presented in the induction videos conflicted with the memories they imagined during the autobiographical recall, which caused confusion. In addition, the videos included photos of people displaying happy and sad expressions. The pre-exposure to emotional faces could have enhanced the sensitivity to this type of stimuli, a priming effect that might have contributed to the mood effects during the emotion judgment task. Study 4 overcame these issues by limiting the emotion induction procedure to viewing sad and happy video scenes portraying animal characters instead of humans. With this procedure, the observed egocentric biases in Study 4 could not be attributed to a pre-exposure to emotional human faces, which strengthens the idea that one's current affective states influence how individuals interpret others' emotional expressions. Moreover, to counteract the transient emotion induction effects, Study 4 continued playing the background music of the videos during the emotion perception blocks. According to the participants' mood ratings, the elicited affective states were overall maintained throughout the emotion judgments blocks, although the intensity of the experienced emotions fainted.

Generally, emotion induction is a complex endeavour, and many different techniques have been used in experimental research to elicit different affective states (Siedlecka & Denson, 2019). In the context of an emotional egocentricity task, the short audio-visual stimuli allowed us to apply within-subject affect manipulations in a single experimental session. Moreover, this technique has been shown to elicit more consistent emotional responses than other methods, such as imagery or autobiographical recall (Zhang et al., 2014). Nevertheless, researchers aiming to use similar affective state manipulations should validate the chosen procedure and materials with the specific population that will compose their study sample, as individual and cultural differences can determine how people react to the affect induction procedure.

6.2.1.1.2 Psychophysical measures of emotion perception biases

The point of subjective equivalence (PSE) was used as the primary dependent variable to assess emotion perception biases. Earlier research had shown that the PSE is a sensitive psychophysical index to measure subtle changes in the perception of facial emotional expressions (Gray et al., 2017; Harris et al., 2016; Marneweck et al., 2013). PSE estimation in traditional psychophysical procedures requires participants to make repeated judgments for each of the morph levels of the happy-sad continuum, leading to a task duration that would have exceeded the transient effects of the emotion induction. To guarantee that emotion judgments would be performed within the emotion manipulation effects, we used an adap-

tive staircase procedure, which reduced the number of trials needed to estimate the PSE. Study 4 further adjusted the adaptive procedure to the participants' baseline PSEs to increase the sensitivity to capture small individual variations in emotion judgments.

The adaptive staircase procedure was easy to implement and successfully reduced the duration of the emotion judgment task to 1 to 3 minutes. However, future studies could also consider more advanced psychophysical adaptive procedures that estimate additional parameters to describe the participants' performance on the emotion perception task (Leek, 2001). For example, in addition to the PSE, some adaptive methods can determine the sensitivity to discriminate between the emotional expressions. This additional measure could provide more nuanced insights into how one's own affective state bias emotion perception (e.g., Qiao-Tasserit et al., 2017).

6.2.1.2 Do emotional egocentric biases reflect general mood effects on cognition?

Our studies contextualised the observed egocentric biases within the phenomenon of affective self-projection. However, mood-congruent biases had been already reported in emotion perception literature. Using different forms of affect manipulations, prior studies showed that individuals tend to be faster in recognising emotional expressions congruent with their current state (Niedenthal et al., 2000, 2001) and judge ambiguous or neutral faces according to their own emotions (Qiao-Tasserit et al., 2017; Schmid & Schmid Mast, 2010). Mood-congruent biases are also indicated by associations between normal variations in mood states and judgments of emotional facial expressions (Jackson & Arlegui-Prieto, 2016).

In addition to influencing emotion perception, mood can affect other cognitive processes. For example, people tend to recall more negative autobiographical memories when they feel sad compared to when feeling neutral or happy (Drace, 2013; Matt et al., 1992). Given the widespread effects of one's mood on how we process information and make decisions (Forgas, 2017), one could argue that egocentric biases in emotion perception are just one more way in which emotional states influence general cognition and behaviour, rather than a specific phenomenon within social cognition.

Nonetheless, the fact that mood-congruent effects may not be limited to socio-emotional processing does not undermine the impact of such biases on social cognition. When inferring how another person is feeling, it is important to keep our emotions apart from the other's (Juckel et al., 2018). Over-attributions of one's own affective experiences, especially when the other person is in a state incongruent with ours, can lead to erroneous assumptions about their affective experiences, motivations and needs. Because understanding others is crucial to navigating in the social world, biased interpretations of others' emotional displays could affect social functioning just as much as dysfunctions in theory of mind or perspective-taking.

6.2.1.3 Potential mechanisms underlying emotional egocentricity

As we have seen, cumulative evidence supports the occurrence of emotional egocentricity, such that one's own affective experiences influence how we judge others' emotional expressions. As a next step, research should investigate the mechanisms that lead to such egocentric biases.

6.2.1.3.1 Mood-guided attention

One way through which one's own emotions could bias the evaluations of others' emotional faces is by guiding our attention toward face regions that are more diagnostic of a particular emotional expression. Studies have shown that some face areas contain more useful information than others for emotion identification (Beaudry et al., 2014; Calvo, Gutiérrez-García, et al., 2018). For example, the mouth region seems to be important to recognise happiness, while the eye region is more relevant to detect sadness and anger (Beaudry et al., 2014; Calvo, Gutiérrez-García, et al., 2018; Wegrzyn et al., 2017). Findings from behavioural studies are supported by eye-tracking data, which has revealed distinct fixation patterns for different emotional expressions (Beaudry et al., 2014; Calvo, Fernández-Martín, et al., 2018; Schurgin et al., 2014).

Due to a limited attentional capacity, we need to distribute our attention as efficiently as possible according to both internal and external demands. As shown by Schurgin et al. (2014), eye movements during emotion perception follow both "stimulus-driven" and "goal-driven" perceptual strategies. Participants fixate their attention to salient features of a stimulus but also attend to regions they think will facilitate recognising a particular emotion (Schurgin et al., 2014). Our expectations about the other's state could, therefore, bias where to allocate our attention. If we use our affective states to guide social inferences, we may fixate our attention more on the face regions that are important for recognising happiness (e.g., mouth region) when we feel happy and focus more on the eye region when we are sad. In other words, by directing our attention to particular face regions, we may be more susceptible to detect expressive signals congruent with our own emotions and neglect other facial cues that reflect a distinct emotion. This hypothesis could be tested in future studies by tracking the participants' eye movements and fixations while performing the emotional egocentricity paradigm developed here.

6.2.1.3.2 Self-other distinction

At the cognitive level, we need to inhibit our own mental and affective states and amplify the representations related to the other to overcome egocentric social inferences. The ability to distinguish and switch between the self- and other-representations has been referred to as self-other distinction or control (Lamm et al., 2016; Sowden & Shah, 2014; Steinbeis, 2016).

Performance in prior emotional egocentricity tasks has been taken as an index of self-other distinction abilities (e.g., Bukowski et al., 2020; Silani et al., 2013; Tomova et al., 2014; von Mohr et al., 2019). In the paradigm by Silani et al. (2013), the participant is simultaneously confronted with multiple perceptual and internal states of the self and the other. These include visual inputs about the type of tactile stimulation being applied to oneself and the target, the perception of the actual tactile stimulation on oneself, and one's own affective reactions to it. In different trials, the participant is asked to report how they or the other person feels in response to the tactile stimulation. To unbiasedly report the own affective states, participants need to inhibit the representations related to the other and enhance the self-representations. To make unbiased inferences about the other, the opposite process should occur: participants need to suppress their own precepts and affective experiences and amplify the representations related to the other. A failure to disengage from the self- or the other-representations could result in egocentric or altercentric emotion judgments, respectively.

In the emotional egocentricity paradigm used in Studies 3 and 4, participants were not confronted with conflicting perceptual information about self and others, nor prompted to switch between judging their own and the other's emotional experience. The only input provided during the emotion judgment task was the target's facial expressions. To overcome egocentric evaluations of the other's emotions, participants had to nevertheless set their current affective states aside and amplify the other's representations. However, self-other distinction in this context was likely more implicit and less deliberate compared to the previous emotional egocentricity paradigms, in which the representations of the self and other were more explicitly confronted.

As outlined by Quesque and Brass (2019), self-other distinction can occur at different levels of information processing, from distinguishing between the self and other bodies (perceptual level) and motor actions (action level) to differentiating between beliefs, intentions and emotions of oneself and others (mental state level). Emotional egocentricity tasks in which participants are asked to switch between reflecting on the self and others' affective experiences (e.g., Silani et al., 2013; Steinbeis & Singer, 2014; von Mohr et al., 2019) fit well within the mental state level described by Quesque and Brass (2019). In contrast, the type of self-other distinction recruited to overcome egocentricity during emotion perception may tap more into the perceptual level, which relates to more implicit and automatic processes.

6.2.1.3.3 Representing the self and the other

A necessary condition for differentiating and switching between self- and other-representations is to represent and correctly "tag" the self and other's states in the first place (Lamm et al., 2016). Individual differences in our capacity to spontaneously represent our own mental and affective states and those of others could therefore determine the degree of egocentricity beyond whether or not self-other distinction is at play.

In Study 3, we found preliminary evidence of an association between egocentric biases and dispositional perspective-taking. Individuals who reported a higher predisposition to take the perspective of others displayed weaker egocentric biases. This supports the idea that a higher tendency to focus on the other's point of view may protect from over-attributing one's own state by spontaneously enhancing the other's representation (Bukowski & Samson, 2017). Nevertheless, given the limited power to test individual differences in Study 3, future research on larger samples is needed to further evaluate the relationship between emotional egocentricity and dispositional perspective-taking.

Parallel to the need to represent others, a lower ability to identify one's own emotional states could also underlie individual differences in emotional egocentricity. If awareness of one's own emotional states is necessary for affective self-projection, one would expect lower egocentric biases in alexithymia, a condition characterised by deficits in recognising one's own emotions (Silani et al., 2008). In this regard, conflicting results have been found. As it would have been hypothesized, one study reported weaker emotional egocentric biases in individuals with higher levels of alexithymia, an effect observed in both a clinical sample with major depressive disorder and control participants (Hoffmann, Banzhaf, et al., 2016). However, no significant associations were detected between emotional egocentric biases and alexithymia in a non-clinical sample using an online adaptation of the classic emotional egocentricity paradigm (von Mohr et al., 2019).

To sum up, the ability to distinguish and shift between self- and other-representations has been considered a central mechanism underlying egocentric biases. However, it is not yet clear how self-other distinction relates to more automatic and implicit processes of affective self-projection. To develop a more complete picture of the role of self-other distinction across different instances of emotional egocentricity, future studies could assess associations between performance in the different emotional egocentricity paradigms, as well as with other tasks thought to involve self-other distinction, such as automatic imitation paradigms (Brass et al., 2009; Spengler et al., 2009). Furthermore, research should not neglect other factors that could also account for individual differences in emotional egocentricity, such as a predisposition to spontaneously represent others' states, or difficulties in recognizing one's own emotions.

6.2.1.4 Limitations of current emotional egocentricity paradigms

In this section, I have discussed and compared two complementary approaches to study self-projection during emotion attributions. While they differ in many aspects, both share some common limitations.

6.2.1.4.1 Inferential accuracy

First, none of the emotional egocentricity paradigms assesses the accuracy of the emotion inferences. By measuring biases in emotion judgments due to within-subject changes in affective states, current paradigms can demonstrate that one's own experiences interfere with emotion inferences. However, the attributed affective states are not compared against the *actual* emotional state experienced by the evaluated targets. Because self-projection is thought to aid social inferences under certain contexts (Ames, 2004b; Tamir & Mitchell, 2013; Todd et al., 2016), it would be interesting to assess whether egocentric tendencies are indeed an effective inferential strategy to understand others' emotions.

Assessing accuracy is much easier in cognitive egocentricity tasks, in which inferences often refer to external objects or agents whose perceptibility can be tracked. For example, in visual perspective-taking paradigms the participants' judgment on what the target sees can be compared to the actual stimuli that are visually accessible to them (Epley, Morewedge, et al., 2004; Samson et al., 2010). Given that feelings and emotions are subjective, estimating the inferential accuracy of affective self-projections is a stronger challenge. Nevertheless, future studies using emotional egocentricity paradigms in which two people are simultaneously tested (e.g., Silani et al., 2008) could achieve this by comparing the participants' inferred emotion ratings against the targets' real ratings.

6.2.1.4.2 Emotion specificity

So far, most emotional egocentricity paradigms have quantified the affective biases by asking participants to make valence or pleasantness ratings on visual analogue scales (Silani et al., 2013; Steinbeis, 2016; von Mohr et al., 2019) or dichotomous emotion judgements (e.g., happy or sad; Study 3 and Study 4). If self-projection entails attributing one's own emotional states to others, we would expect not only congruency in the valence of the experienced and inferred state, but also specificity in the type of emotion. For example, feeling angry should specifically increase the tendency to attribute anger to others, but not other negative emotions such as sadness. To assess the specificity in emotional egocentricity, paradigms could broaden the range of emotional experiences induced and judged by participants.

6.2.1.4.3 Ecological validity

Finally, another limitation is the reduced ecological validity of the paradigms. In Studies 3 and 4, participants were asked to make emotion judgments of morphed facial expressions of only one or two identities, presented very briefly and devoid of context. This procedure allowed to control for potential confounding variables and facilitated the detection of perceptual biases. However, such impoverished emotional displays are far from the type of information we encounter in daily situations. We typically have more time to interpret the other's emotional expressions and integrate further contextual cues, such as other overt behaviour (e.g., gestures, postures, tone of voice) or information about the events that preceded the other's emotional display, all of which can condition how emotional expressions are perceived (Barrett et al., 2011; Gray et al., 2017; Juckel et al., 2018).

Research should move towards more naturalistic approaches in which many of these elements (e.g., emotional expressions, gestures, situational information) are combined. Towards this aim, our research group has developed an emotional egocentricity paradigm that integrated both facial emotional expressions and information about the situational context (Weigand et al., 2021). In particular, participants were asked to judge facial emotional expressions while being previously confronted with congruent or incongruent emotion-inducing situations. Results showed that how participants evaluated the situational context egocentrically biased their judgements about the observed emotional expressions.

Emotional egocentricity paradigms that integrate multiple cues will make causal interpretations more challenging, but the effects will likely be more generalisable to daily life situations. This will also allow testing the extent to which affective self-projection is recruited when individuals are in contextually rich environments. Predictively, self-projection may primarily be used when we have limited information about the other person and their context or when their emotional displays are ambiguous. When more information is available, individuals may rely on alternative inferential strategies that could be more effective in understanding what the other is experiencing.

6.2.2 Self-projection to similar and dissimilar others

According to the similarity contingency model, the degree to which people recruit their own mental states during social inferences depends on their perceived similarity with the other person (Ames, 2004b). Using a measure of cognitive self-projection, Study 4 replicated the perceived similarity effect found in previous studies. Participants' own ratings significantly predicted the attitudes estimated for the similar targets but not for the dissimilar targets. This result adds to the extensive bulk of research showing stronger attribution of one's own traits, beliefs and attitudes to people we perceive as similar to oneself compared to those who are dissimilar (Ames, 2004b, 2004a; Clement & Krueger, 2002; Davis, 2017; Tamir & Mitchell, 2013; Woo & Mitchell, 2020).

However, results from Study 4 did not support the similarity contingency model for affective self-projection. Against our main hypothesis, the degree to which one's own affective states biased emotion judgments was not conditional on the similarity with the target. Our data was more indicative that egocentric biases during emotion judgments were comparable for similar and dissimilar targets. Differences in the type of inferential process, as well as characteristics of the similarity manipulation performed, could account for the discrepancy in the modulation of cognitive and affective self-projection.

6.2.2.1 Cognitive vs affective inferences

Perhaps the most apparent difference between cognitive and affective self-projection tasks is the type of information participants needed to infer. In the cognitive self-projection task in Study 4, participants were asked to predict attitudes and preferences of similar and dissimilar targets. In contrast, the emotional egocentricity task assessed how participants evaluated others' emotional facial expressions. Comparing the results of both tasks, one could conclude that perceived similarity plays a role in self-projection in the cognitive domain but not in the affective domain. However, there is evidence that individuals attribute their feelings of thirst and coldness to similar targets but not to dissimilar targets (O'Brien & Ellsworth, 2012). Given that visceral feelings and sensations tap more into the affective domain, it is unlikely that the type of attribute inferred entirely accounts for the null effect of perceived similarity on emotional egocentricity found in Study 4.

6.2.2.2 High- vs low-level inferential processes

A more likely explanation of the discrepancy between the impact of self-projection between the two tasks may lie in the nature of the inferential process assessed. Most cognitive self-projection paradigms, including the one used here, are assumed to recruit higher-order processes of social cognition, such as theory of mind or perspective-taking (Quesque & Rossetti, 2020). Higher-level processes of affective inferences were also likely recruited in O'Brien and Ellsworth (2012), as participants had to predict the target's visceral states from a written story that described the event the target was experiencing. In contrast, reading other's emotional facial expressions tends to be characterised as a lower-level and more automatic process of emotion attribution (Etchepare & Prouteau, 2018; Quesque & Rossetti, 2020; Tracy & Robins, 2008).

Even though emotion perception can be top-down modulated by contextual factors (Barrett et al., 2011), the impact of social cues may be weaker on such automatic and lower-level processes compared to more evaluative forms of emotion attribution. That is, emotional egocentricity may be modulated by similarity but to a lesser degree than cognitive egocentricity. If this were the case, our study might not have had enough power to detect such small effects,

as our effect size of interest (Cohen's d = 0.35) was based on meta-analytic estimates of the influence of social categorisation on cognitive self-projection (Robbins & Krueger, 2005).

In addition to power issues, our experimental design may have limited the chance for the participants to consider contextual factors during the emotion judgements. Variables such as the gender, race or gaze direction of a face have been shown to bias the evaluation of emotional expressions, even when presented at comparably short times as in Study 4 (Adams Jr. & Kleck, 2005; Craig & Lipp, 2018; Harris et al., 2016). However, implicit associations with factors related to the target's identity may be more readily available than evaluations of the perceived similarity with the other. Given the quick responses required in the emotional egocentricity task, participants may not have had enough time to retrieve and integrate the previously learnt information about the target. This contrasts with the cognitive self-projection task, which participants did not have strict time constraints to make their judgements about the targets' attitudes and preferences.

6.2.2.3 Perceived similarity: a context-dependent evaluation?

Finally, how similarity was determined may have also conditioned the observed pattern of results. In Study 4, perceived similarity was manipulated by assigning the same or a different cognitive style to the participants and the targets they evaluated. The cognitive style ostensibly reflected how numerical information was processed, a characteristic that was unrelated to the content of the inferences made during the cognitive and affective self-projection tasks. Laboratory manipulations based on arbitrary features can be powerful tools to create a sense of similarity without being confounded by pre-existing stereotypes on the social groups the target belongs to (Clement & Krueger, 2002). However, one limitation of this kind of manipulations is that it could create a feeling of (dis)similarity that may not be relevant for the type of social inferences being investigated.

In natural social interactions, similarity is evaluated according to the current setting and may even be founded on factors that might not reflect the "true" similarity (Ames, 2004b). For example, we may feel similar to a person in the context of an election because we share the same political attitudes; but evaluate the same person as very dissimilar when we are at work because we have different professional roles. Depending on what we need to infer and in which context, we may recruit more or less self-projection according to the situational evaluation of similarity.

It could be hypothesised that knowing whether a target processes cognitive information in a similar way as oneself may not inform if the person will show similar emotional reactions. If this is the case, participants might not have recruited this information about the targets during the emotional egocentricity task. To test this hypothesis, future studies could manipulate perceptions of similarity based on emotion-related vs non-emotion attributes and

examine whether such manipulations differentially modulate affective vs cognitive self-projection.

In sum, data from Study 4 suggested that perceived similarity does not modulate the degree of egocentric biases when reading others' emotional expressions, at least not to the same extent as it influences more cognitive or evaluative forms of social inferences. Importantly, these results should not be interpreted as conclusive proof that similarity does not play a role in affective self-projection. Further data is required to assess whether this negative finding is related to the type of inferences studied, the similarity manipulation, power issues, or if it reflects a true null effect of similarity on affective self-projections.

6.3 Self-other control: a common mechanism behind altercentric and egocentric processes?

As discussed earlier, self-other distinction has been hypothesised as one mechanism underlying (emotional) egocentricity. Notably, the ability to switch between representations of self and others also plays a central role in altercentric phenomena (Sowden & Shah, 2014). Self-other distinction and control has been proposed as a basic neurocognitive mechanism across socio-cognitive functions (Happé et al., 2017; Sowden & Shah, 2014). In some cases, the self-representation needs to be suppressed (e.g., when overcoming self-projection during social inferences), while other situations require one to distance oneself from the representations of others (e.g., when counteracting automatic imitation tendencies).

The idea self-other distinction subserves multiple socio-cognitive functions is supported by studies showing associations between the control of automatic imitation and performance in other social cognition domains, such as perspective-taking and theory of mind (Spengler, Bird, et al., 2010; Spengler, Cramon, et al., 2010). In addition, training participants to inhibit automatic imitation has shown to induce changes in their perspective-taking abilities (Santiesteban, White, et al., 2012) and empathic responses to pain (de Guzman et al., 2016), which bring more direct and causal evidence for this hypothesis.

On the neural level, the right TPJ has been consistently implicated in tasks requiring a shift from self to other representations (Martin et al., 2018; Santiesteban, Banissy, et al., 2012; Spengler et al., 2009), and vice versa (Brass et al., 2005; Hogeveen et al., 2015; Santiesteban, Banissy, et al., 2012; Sowden & Catmur, 2015; Spengler et al., 2009). On the affective domain, self-other distinction is thought to rely on a region located anterior to the right TPJ, namely the supramarginal gyrus (Bukowski et al., 2020; Silani et al., 2013).

Though the notion of a common underlying mechanism is appealing, at present, it remains unclear how self-other distinction unfolds across different levels of social information processing (e.g., perceptual, action and mental state; Quesque & Brass, 2019) and domains (e.g., affective vs cognitive mental states). According to personal goals and demands of the social

context, individuals need to inhibit or amplify different co-activated representations of self and others. One challenge for future research will be to relate and study the interplay of self-other control abilities across altercentric and egocentric processes.

6.4 Implications and future directions

In this dissertation, I have reviewed extensive literature assessing contextual determinants of social cognition. Many of the studies, including ours, tested the effects of a single contextual variable, or two at most. However, what defines a social context is rarely one isolated factor. It is the interplay between multiple social and contextual cues that gives meaning to a situation and allows one to understand other agents. A social cue that may be perceived as a sign of social interest in one setting (e.g., direct gaze with a happy expression) could acquire a completely different meaning when displayed in another situation (e.g., direct gaze with an angry expression). Even when situational factors are not explicitly manipulated, contextual information may be implicitly provided by the particular attributes of the agent presented (Hess & Fischer, 2013). Furthermore, as seen in Part II of this dissertation, even a person's internal states may influence how social cues are perceived and interpreted.

Given that context is inherent in social interactions, research on altercentric and egocentric socio-cognitive phenomena would benefit from analysing contextual cues in a more integrated manner. Instead of focusing on single isolated variables and treating other contextual factors as "confounders", it may be more fruitful to acknowledge and exploit the richness of a social environment, even in controlled laboratory settings. The increasing adoption of second-person neuroscience paradigms manifests the trend towards more ecologically valid research (Schilbach et al., 2013). By developing experimental settings that simulate better a real-life interaction, second-person neuroscience studies have shown that some of the neurocognitive mechanisms that had been identified with tasks relying on social observation differ from the brain networks recruited when individuals are actually engaged in social interactions (Redcay & Schilbach, 2019).

Adopting a more situated notion of social cognition may also help elucidate potential mechanisms underlying disorders of social interaction. One benefit of examining socio-cognitive processes together with the context in which they unfold is that it encourages researchers to specify the situations in which a particular phenomenon is manifested, as well as its boundaries. By using this integrated approach, research may be able to determine not only what socio-cognitive processes are necessary for social functioning, but also in which contexts those become relevant or should be inhibited. This could be specially important to understand socio-cognitive deficits in ASC. As previously reviewed, there is evidence suggesting that atypical social functioning in individuals with ASC may arise from a reduced ability to implicitly encode and integrate contextual cues to regulate their behaviour and cognition (Baez & Ibanez, 2014; Wang & Hamilton, 2012). By linking the social cognition

deficits in clinical samples to the context in which they are manifested, future studies may be able to identify more fine-grained domains of disability, but also detect preserved capacities that could be enhanced.

6.5 A note on publishing non-significant results

Throughout this dissertation, I have presented and discussed more non-significant results than I would have expected at the beginning of my doctoral studies. Under the framework of null hypothesis significance testing (NHST), non-significant effects are a possible outcome. However, the trends in psychological science in the last decades have not promoted the publication of negative findings (Joober et al., 2012). Publication biases have been long documented, whereby positive results are more prone to be considered for publication, more favourably reviewed and more likely cited. This is also true for fields such as experimental and social psychology (Earp & Trafimow, 2015; Francis, 2012), in which non-replications and null results are likely due to the high contextual variability of social human behaviour (Van Bavel et al., 2016).

A publication bias is problematic as it can give the perception that a false effect is true and can overestimate the size of true effects (Francis, 2012). A complex interplay between academic pressures and incentives, and the higher publication rate of positive findings, have stimulated questionable research practices (John et al., 2012). Selective reporting of measures and results, p-hacking, hypothesising after the results are known (HARKing), and data-peeking are some practices that increase the probability of false positives and reduce the interpretability of the published statistical results. Given that such practices are frequent in psychology research (John et al., 2012), it is expected that many of the positive reported effects are false (Ioannidis, 2005). In contrast, negative results that are more likely to be true-negative results are not reported.

6.5.1 Different types of non-significant results

On their own, non-significant outcomes in NHST do not allow conclusions about the presence or absence of the effect. A non-significant result does not necessarily provide evidence against the hypothesised effects as it could also indicate data insensitivity. Joober et al. (2012) defined three different categories of negative results. Conclusive negative results refer to those that bring clear evidence of the absence of an effect or the opposite effect as originally hypothesised. Explorative negative results are well-designed and adequately powered studies that yield neutral or opposite results based on exploratory data analysis. Finally, inconclusive negative results are those in which there is no evidence of an effect in a study that was inadequately powered.

Different statistical approaches have been recently proposed as complementary strategies to NHST to make null results informative (Dienes, 2014; Harms & Lakens, 2018; Lakens et al., 2018). On the one hand, equivalence testing is a tool based on frequentist statistics that allows rejecting the presence of an effect that is as large or larger than the smallest effect size of interest (Lakens et al., 2018). On the other hand, Bayesian statistics propose the use of Bayes factors, which allow estimating the relative evidence for the null hypothesis over the alternative hypothesis given the data (Dienes, 2014).

These statistical tools helped us interpret and assess the implications of the non-significant results obtained in the work presented here. For example, we demonstrated that the non-significant effect of similarity on emotional egocentricity in Study 4 reflected a conclusive negative result. In other words, we could provide evidence that similarity in our study did not modulate the degree of egocentric biases during quick emotion judgements, at least not to the extent that it did on cognitive self-projection. In contrast, equivalence testing in Study 1 revealed that the study lacked statistical power to detect the hypothesised intranasal oxytocin effects on mimicry. Study 1, therefore, would fit under the category of inconclusive negative results, according to Joober et al. (2012).

6.5.2 Transparent reporting and open science

Reporting non-significant effects, even when these are inconclusive, is crucial to assess the overall significance of a research field. A single null finding may not be meaningful on its own. However, cumulative negative findings can change the confidence in the original results and contribute to a more accurate estimation of the true effect size in meta-analyses (Earp & Trafimow, 2015; Joober et al., 2012).

Some of the null results presented here have made the way to this dissertation a bit more challenging. However, it has also pushed me to be more critical with published literature and adopt research practices that promote reproducibility and transparency. Whenever possible, we preregistered the hypotheses, methods and statistical analyses of the experiments (Study 2 and Study 4) and shared the collected data and statistical analyses in public repositories (Studies 1-4). Moreover, manuscripts have been published in open-access formats, and some were available as preprints beforehand. While there is still a lot to improve, and not all best practices can be adopted for all research projects, each small step towards open and reproducible science will hopefully help achieve a more credible and realistic picture of research findings.

6.6 Conclusion

Social cognition involves complex mechanisms to understand and navigate through social interactions. In this dissertation, I have presented some evidence supporting the idea that social-cognitive processes are situated and depend on multiple contextual factors, including the person's own state of mind, their perceptions about the other, their goals, and interpersonal relationship. Due to the constant and complex interchange between oneself and others, social functioning requires a balance between being influenced by, representing, and switching between our own mental and affective states and those of others.

The studies discussed in this dissertation have also demonstrated the complexity of investigating highly contextual phenomena. Extending findings from earlier studies (e.g., investigating modulating factors of a previously reported effect) was initially a primary motivator of my research. However, the failure to (conceptually) replicate several of the previously reported phenomena highlights the need to revisit some of the effects and systematically re-evaluate the constructs and operationalisations social cognition research aims to investigate. Studying such processes isolated or without taking the context in which they unfold into consideration may give us just one small part of the big picture of social cognition.

Glossary of terminology

Altercentrism	Influence of the representations of beliefs, actions, perceptions, or emotions of others on the representations related to the self.
Automatic imitation	Form of stimulus-response compatibility whereby observation of another's action prompts the tendency to produce an identical action.
Egocentric bias	Inaccurate social inference due to an over-attribution of one's own mental states onto others.
Egocentrism	Influence of one's own beliefs, actions, perceptions or emotions on the representations related to the other.
Emotion perception	Ability to determine the affective state of another person by interpreting their physical changes, such as facial expressions, body postures and prosody.
Empathic concern	Tendency to react with feelings of sympathy and concern for unfortunate others.
Empathy	Ability to share and understand the emotional states of another person, with the recognition that the other is the source of one's state.
Imitative compatibility	Tendency to perform topographically similar movements or body configurations to an observed action.
Mimicry	Spontaneous imitation of another person's behaviours or emotional expressions that occurs without conscious awareness.

Oxytocin	Hormone and neuromodulator that plays a role in both social and non-social behaviour by maintaining stability through changing environments.
Perspective-taking	Ability to consider a situation from the point of view of another agent.
Self-other control	Ability to switch between coactivated representations related to the self and to others.
Self-other distinction	Ability to distinguish between the representations of our own beliefs, actions, perceptions, sensations and emotions, and those of others.
Self-projection	Inferential act of attributing one's own mental states and behaviours onto others.
Social cognition	Conjunction of cognitive processes that allow us to understand other agents and to coordinate with the social world.
Social context	Specific setting in which a cognitive process, social interaction or action occurs or unfolds. The social context can include factors related to the situation, personal characteristics and the relationship between the interaction agents.
Social inference	Judgment about the mental states of another person or group.
Spatial compatibility	Tendency to respond in the same relative spatial position as the observed movement.
Theory of mind	Ability to infer the beliefs, intentions and emotions of others.

List of figures

Study 1

Study 2

Study 3

Study 4

List of tables

Study 1: Facial mimicry, reward & oxytocin

Study 2: Automatic imitation, gaze & autistic traits

Study 3: Emotional egocentricity & socio-emotional traits

Study 4: Emotional egocentricity & similarity

References

Adams Jr., R. B., & Kleck, R. E. (2005). Effects of Direct and Averted Gaze on the Perception of Facially Communicated Emotion. *Emotion*, *5*(1), 3–11. https://doi.org/10.1037/1528-3542.5.1.3

Aguado, L., Martínez-García, N., Solís-Olce, A., Dieguez-Risco, T., & Hinojosa, J. A. (2018). Effects of affective and emotional congruency on facial expression processing under different task demands. *Acta Psychologica*, *187*, 66–76. https://doi.org/10.1016/j.actpsy.2018.04.013

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Ames, D. R. (2004a). Inside the Mind Reader's Tool Kit: Projection and Stereotyping in Mental State Inference. *Journal of Personality and Social Psychology*, *87*(3), 340–353. https://doi.org/10.1037/0022-3514.87.3.340

Ames, D. R. (2004b). Strategies for Social Inference: A Similarity Contingency Model of Projection and Stereotyping in Attribute Prevalence Estimates. *Journal of Personality and Social Psychology*, *87*(5), 573–585. https://doi.org/10.1037/0022-3514.87.5.573

Ames, D. R., Weber, E. U., & Zou, X. (2012). Mind-reading in strategic interaction: The impact of perceived similarity on projection and stereotyping. *Organizational Behavior and Human Decision Processes*, *117*(1), 96–110. https://doi.org/10.1016/j.obhdp.2011.07.007

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker: *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1080/17470210903281582

Baez, S., & Ibanez, A. (2014). The effects of context processing on social cognition impairments in adults with Asperger's syndrome. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00270

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism,

males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. https://doi.org/10.1023/a:1005653411471

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in Emotion Perception. *Current Directions in Psychological Science*, *20*(5), 286–290. https://doi.org/10.1177/0963721411422522

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bazinger, C., & Kühberger, A. (2012). Is social projection based on simulation or theory? Why new methods are needed for differentiating. *New Ideas in Psychology*, *30*(3), 328–335. https://doi.org/10.1016/j.newideapsych.2012.01.002

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition and Emotion*, *28*(3), 416–432. https://doi.org/10.1080/02699931.2013.833500

Bernier, R., Dawson, G., Webb, S., & Murias, M. (2007). EEG mu rhythm and imitation impairments in individuals with autism spectrum disorder. *Brain and Cognition*, *64*(3), 228–237. https://doi.org/10.1016/j.bandc.2007.03.004

Bethlehem, R. A. I., Baron-Cohen, S., van Honk, J., Auyeung, B., & Bos, P. A. (2014). The oxytocin paradox. *Frontiers in Behavioral Neuroscience*, *8*. https://doi.org/10.3389/fnbeh.2014.00048

Bird, G., & Viding, E. (2014). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, *47*, 520–532. https://doi.org/10.1016/j.neubiorev.2014.09.021

Blocker, H. S., & McIntosh, D. N. (2016). Automaticity of the interpersonal attitude effect on facial mimicry: It takes effort to smile at neutral others but not those we like. *Motivation and Emotion*, *40*(6), 914–922. https://doi.org/10.1007/s11031-016-9581-7

Bortoletto, M., Mattingley, J. B., & Cunnington, R. (2013). Effects of Context on Visuomotor Interference Depends on the Perspective of Observed Actions. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0053248

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, *77*(3), 343–352. https://doi.org/10.1016/j.biopsycho.2007.11.008

Boyer, T. W., Longo, M. R., & Bertenthal, B. I. (2012). Is automatic imitation a specialized form of stimulus-response compatibility? Dissociating imitative and spatial compatibilities. *Acta Psychologica*, *139*(3), 440–448. https://doi.org/10.1016/j.actpsy.2012.01.003

Bradford, E. E. F., Hukker, V., Smith, L., & Ferguson, H. J. (2018). Belief-attribution in adults with and without autistic spectrum disorders. *Autism Research*, *11*(11), 1542–1553. https://doi.org/10.1002/aur.2032

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van 't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*, 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between Observed and Executed Finger Movements: Comparing Symbolic, Spatial, and Imitative Cues. *Brain and Cognition*, *44*(2), 124–143. https://doi.org/10.1006/brcg.2000.1225

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia*, *43*(1), 89–98. https://doi.org/10.1016/j.neuropsychologia.2004.06.018

Brass, M., & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495. https://doi.org/10.1016/j.tics.2005.08.007

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2359–2367. https://doi.org/10.1098/rstb.2009.0066

Brunyé, T. T., Ditman, T., Giles, G. E., Mahoney, C. R., Kessler, K., & Taylor, H. A. (2012). Gender and autistic personality traits predict perspective-taking ability in typical adults. *Personality and Individual Differences*, *52*(1), 84–88. https://doi.org/10.1016/j.paid.2011.09.004

Bukowski, H., & Samson, D. (2017). New Insights into the Inter-Individual Variability in Perspective Taking. *Vision*, *1*(1), 8. https://doi.org/10.3390/vision1010008

Bukowski, H., Tik, M., Silani, G., Ruff, C. C., Windischberger, C., & Lamm, C. (2020). When differences matter: RTMS/fMRI reveals how differences in dispositional empathy translate to distinct neural underpinnings of self-other distinction in empathy. *Cortex*, *128*, 143–161. https://doi.org/10.1016/j.cortex.2020.03.009

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Calvo, M. G., Fernández-Martín, A., Gutiérrez-García, A., & Lundqvist, D. (2018). Selective eye fixations on diagnostic face regions of dynamic emotional expressions: KDEF-dyn database. *Scientific Reports*, *8*(1), 17039. https://doi.org/10.1038/s41598-018-35259-w

Calvo, M. G., Gutiérrez-García, A., & Del Líbano, M. (2018). What makes a smiling face look happy? Visual saliency, distinctiveness, and affect. *Psychological Research*, *82*(2), 296–309. https://doi.org/10.1007/s00426-016-0829-3

Campbell, M. E. J., & Cunnington, R. (2017). More than an imitation game: Top-down modulation of the human mirror system. *Neuroscience & Biobehavioral Reviews*, *75*, 195–202. https://doi.org/10.1016/j.neubiorev.2017.01.035

Cath, D. C., Ran, N., Smit, J. H., Balkom, A. J. L. M. van, & Comijs, H. C. (2008). Symptom Overlap between Autism Spectrum Disorder, Generalized Social Anxiety Disorder and Obsessive-Compulsive Disorder in Adults: A Preliminary Case-Controlled Study. *Psychopathology*, *41*(2), 101–110. https://doi.org/10.1159/000111555

Catmur, C., & Heyes, C. (2011). Time course analyses confirm independence of imitative and spatial compatibility. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(2), 409–421. https://doi.org/10.1037/a0019325

Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: The role of experience in the development of imitation and the mirror system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2369–2380. https://doi.org/10.1098/rstb.2009.0048

Cavanagh, J., & Geisler, M. W. (2006). Mood effects on the ERP processing of emotional intensity in faces: A P3 investigation with depressed students. *International Journal of Psychophysiology*, *60*(1), 27–33. https://doi.org/10.1016/j.ijpsycho.2005.04.005

Champely, S. (2018). *pwr: Basic Functions for Power Analysis* (R package version 1.2-2) [Computer software]. https://CRAN.R-project.org/package=pwr

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Chartrand, T. L., & Lakin, J. L. (2013). The Antecedents and Consequences of Human Behavioral Mimicry. *Annual Review of Psychology*, *64*(1), 285–308. https://doi.org/10.1146/annurev-psych-113011-143754

Chartrand, T. L., & van Baaren, R. (2009). Human mimicry. In *Advances in experimental social psychology, Vol 41* (pp. 219–274). Elsevier Academic Press. https://doi.org/10.1016/S0065-2601(08)00405-X

Chatel-Goldman, J., Schwartz, J.-L., Jutten, C., & Congedo, M. (2013). Non-local mind from the perspective of social cognition. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00107

Cheung, E. O., Slotter, E. B., & Gardner, W. L. (2015). Are you feeling what I'm feeling? The role of facial mimicry in facilitating reconnection following social exclusion. *Motivation and Emotion*, *39*(4), 613–630. https://doi.org/10.1007/s11031-015-9479-9

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The Social Motivation Theory of Autism. *Trends in Cognitive Sciences*, *16*(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Clark-Elford, R., Nathan, P. J., Auyeung, B., Voon, V., Sule, A., Müller, U., Dudas, R., Sahakian, B. J., Phan, K. L., & Baron-Cohen, S. (2014). The effects of oxytocin on social reward learning in humans. *The International Journal of Neuropsychopharmacology*, *17*(2), 199–209. https://doi.org/10.1017/S1461145713001120

Clement, R. W., & Krueger, J. (2002). Social Categorization Moderates Social Projection. *Journal of Experimental Social Psychology*, *38*(3), 219–231. https://doi.org/10.1006/jesp.2001.1503

Coburn, P. I., Bernstein, D. M., & Begeer, S. (2015). A new paper and pencil task reveals adult false belief reasoning bias. *Psychological Research*, *79*(5), 739–749. https://doi.org/10.1007/s00426-014-0606-0

Collova, J. R., Kloth, N., Crookes, K., Burton, N., Chan, C. Y. H., Hsiao, J. H., & Rhodes, G. (2017). A new other-race effect for gaze perception. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(11), 1857–1863. https://doi.org/10.1037/xhp0000460

Cook, J. L., & Bird, G. (2012). Atypical social modulation of imitation in autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *42*(6), 1045–1051. https://doi.org/10.1007/s10803-011-1341-7

Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., De Coster, L., Radkova, I., Deschrijver, E., & Brass, M. (2018). Automatic imitation: A meta-analysis. *Psychological Bulletin*, *144*(5), 453–500. https://doi.org/10.1037/bul0000143

Cracco, E., Genschow, O., Radkova, I., & Brass, M. (2018). Automatic imitation of pro- and antisocial gestures: Is implicit social behavior censored? *Cognition*, *170*, 179–189. https://doi.org/10.1016/j.cognition.2017.09.019

Craig, B. M., & Lipp, O. V. (2018). Facial age cues and emotional expression interact asymmetrically: Age cues moderate emotion categorisation. *Cognition and Emotion*, *32*(2), 350–362. https://doi.org/10.1080/02699931.2017.1310087

Critcher, C. R., & Dunning, D. (2009). Egocentric pattern projection: How implicit personality theories recapitulate the geography of the self. *Journal of Personality and Social Psychology*, *97*(1), 1–16. https://doi.org/10.1037/a0015670

Dalmaso, M., Galfano, G., & Castelli, L. (2015). The Impact of Same- and Other-Race Gaze Distractors on the Control of Saccadic Eye Movements. *Perception*, *44*(8–9), 1020–1028. https://doi.org/10.1177/0301006615594936

Dalmaso, M., Pavan, G., Castelli, L., & Galfano, G. (2012). Social status gates social attention in humans. *Biology Letters*, *8*(3), 450–452. https://doi.org/10.1098/rsbl.2011.0881

Darda, K. M., Butler, E. E., & Ramsey, R. (2020). Individual differences in social and non-social cognitive control. *Cognition*, 104317. https://doi.org/10.1016/j.cognition.2020.104317

Darda, K. M., & Ramsey, R. (2019). The inhibition of automatic imitation: A meta-analysis and synthesis of fMRI studies. *NeuroImage*, *197*, 320–329. https://doi.org/10.1016/j.neuroimage.2019.04.059

Davis, M. H. (1980). A Multidimensional Approach to Individual Differences in Empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85.

Davis, M. H. (2017). Social projection to liked and disliked targets: The role of perceived similarity. *Journal of Experimental Social Psychology*, *70*, 286–293. https://doi.org/10.1016/j.jesp.2016.11.012

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology*, *70*(4), 713–726. https://doi.org/10.1037/0022-3514.70.4.713

De Coster, L., Mueller, S. C., T'Sjoen, G., De Saedeleer, L., & Brass, M. (2014). The influence of Oxytocin on automatic motor simulation. *Psychoneuroendocrinology*, *50*, 220–226. https://doi.org/10.1016/j.psyneuen.2014.08.021

de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions: Emotional body perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 149–158. https://doi.org/10.1002/wcs.1335

de Guzman, M., Bird, G., Banissy, M. J., & Catmur, C. (2016). Self–other control processes in social cognition: From imitation to empathy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150079. https://doi.org/10.1098/rstb.2015.0079

Deaner, R. O., Shepherd, S. V., & Platt, M. L. (2007). Familiarity accentuates gaze cuing in women but not men. *Biology Letters*, *3*(1), 64–67. https://doi.org/10.1098/rsbl.2006.0564

Deng, H., & Hu, P. (2018). Matching Your Face or Appraising the Situation: Two Paths to Emotional Contagion. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02278

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00781

Dimberg, U., Andréasson, P., & Thunberg, M. (2011). Emotional Empathy and Facial Reactions to Facial Expressions. *Journal of Psychophysiology*, *25*(1), 26–31. https://doi.org/10.1027/0269-8803/a000029

Domes, G., Marx, L., Spenthof, I., & Heinrichs, M. (2016). The German Version of the Gaze Anxiety Rating Scale (GARS): Reliability and Validity. *PloS One*, *11*(3), e0150807. https://doi.org/10.1371/journal.pone.0150807

Drace, S. (2013). Evidence for the role of affect in mood congruent recall of autobiographic memories. *Motivation and Emotion*, *37*(3), 623–628. https://doi.org/10.1007/s11031-012-9322-5

Drimalla, H., Landwehr, N., Hess, U., & Dziobek, I. (2019). From face to face: The contribution of facial mimicry to cognitive and emotional empathy. *Cognition & Emotion*, *33*(8), 1672–1686. https://doi.org/10.1080/02699931.2019.1596068

Drimalla, H., Scheffer, T., Landwehr, N., Baskow, I., Roepke, S., Behnia, B., & Dziobek, I. (2020). Towards the automatic detection of social biomarkers in autism spectrum disorder: Introducing the simulated interaction task (SIT). *Npj Digital Medicine*, *3*(1), 1–10. https://doi.org/10.1038/s41746-020-0227-5

Duffy, K. A., & Chartrand, T. L. (2015). Mimicry: Causes and consequences. *Current Opinion in Behavioral Sciences*, *3*, 112–116. https://doi.org/10.1016/j.cobeha.2015.03.002

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, *38*(3), 464–473.

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00621

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*(1), 351–362. https://doi.org/10.3758/BRM.42.1.351

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective Taking as Egocentric Anchoring and Adjustment. *Journal of Personality and Social Psychology*, *87*(3), 327–339. https://doi.org/10.1037/0022-3514.87.3.327

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768. https://doi.org/10.1016/j.jesp.2004.02.002

Etchepare, A., & Prouteau, A. (2018). Toward a Two-Dimensional Model of Social Cognition in Clinical Neuropsychology: A Systematic Review of Factor Structure Studies. *Journal of the International Neuropsychological Society: JINS*, *24*(4), 391–404. https://doi.org/10.1017/S1355617717001163

Farmer, H., Mahmood, R., Gregory, S. E. A., Tishina, P., & Hamilton, A. F. de C. (2021). Dynamic emotional expressions do not modulate responses to gestures. *Acta Psychologica*, *212*, 103226. https://doi.org/10.1016/j.actpsy.2020.103226

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feldman, R., Monakhov, M., Pratt, M., & Ebstein, R. P. (2016). Oxytocin Pathway Genes: Evolutionary Ancient System Impacting on Human Affiliation, Sociality, and Psychopathology. *Biological Psychiatry*, *79*(3), 174–184. https://doi.org/10.1016/j.biopsych.2015.08.008

FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour*, *3*(5), 426–435. https://doi.org/10.1038/s41562-019-0590-x

Fischer, A., & Hess, U. (2017). Mimicking emotions. *Current Opinion in Psychology*, *17*, 151–155. https://doi.org/10.1016/j.copsyc.2017.07.008

Forbes, P. A. G., Korb, S., Radloff, A., & Lamm, C. (2021). The effects of self-relevance vs. Reward value on facial mimicry. *Acta Psychologica*, *212*, 103193. https://doi.org/10.1016/j.actpsy.2020.103193

Forbes, P. A. G., Pan, X., & de C. Hamilton, A. F. (2016). Reduced Mimicry to Virtual Reality Avatars in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *46*(12), 3788–3797. https://doi.org/10.1007/s10803-016-2930-2

Forbes, P. A. G., Wang, Y., & de C. Hamilton, A. F. (2017). STORMy Interactions: Gaze and the Modulation of Mimicry in Adults on the Autism Spectrum. *Psychonomic Bulletin & Review*, *24*(2), 529–535. https://doi.org/10.3758/s13423-016-1136-0

Forgas, J. P. (2017). Mood Effects on Cognition: Affective Influences on the Content and Process of Information Processing and Behavior. In *Emotions and Affect in Human Factors and Human-Computer Interaction* (pp. 89–122). Elsevier. https://doi.org/10.1016/B978-0-12-801851-4.00003-3

Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review*, *19*(6), 975–991. https://doi.org/10.3758/s13423-012-0322-y

Freitag, C. M., Retz-Junginger, P., Retz, W., Seitz, C., Palmason, H., Meyer, J., Rösler, M., & Gontard, A. von. (2007). Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ)—Die Kurzversion AQ-k. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*. https://econtent.hogrefe.com/doi/abs/10.1026/1616-3443.36.4.280

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze Cueing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin*, *133*(4), 694–724. https://doi.org/10.1037/0033-2909.133.4.694

Genschow, O., & Schindler, S. (2016). The influence of group membership on cross-contextual imitation. *Psychonomic Bulletin & Review*, *23*(4), 1257–1265. https://doi.org/10.3758/s13423-015-0983-4

Genschow, O., van Den Bossche, S., Cracco, E., Bardi, L., Rigoni, D., & Brass, M. (2017). Mimicry and automatic imitation are not correlated. *PLOS ONE*, *12*(9), e0183784. https://doi.org/10.1371/journal.pone.0183784

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'Reading the Mind in the Voice' Test-Revised: A Study of Complex Emotion Recognition in Adults with and Without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, *37*(6), 1096–1106. https://doi.org/10.1007/s10803-006-0252-5

Gray, K. L. H., Barber, L., Murphy, J., & Cook, R. (2017). Social interaction contexts bias the perceived expressions of interactants. *Emotion*, *17*(4), 567–571. https://doi.org/10.1037/emo0000257

Guastella, A. J., Hickie, I. B., McGuinness, M. M., Otis, M., Woods, E. A., Disinger, H. M., Chan, H.-K., Chen, T. F., & Banati, R. B. (2013). Recommendations for the standardisation of oxytocin nasal administration and guidelines for its reporting in human research. *Psychoneuroendocrinology*, *38*(5), 612–625. https://doi.org/10.1016/j.psyneuen.2012.11.019

Guéguen, N., & Martin, A. (2009). Incidental Similarity Facilitates Behavioral Mimicry. *Social Psychology*, *40*(2), 88–92. https://doi.org/10.1027/1864-9335.40.2.88

Haffey, A., Press, C., O'Connell, G., & Chakrabarti, B. (2013). Autistic traits modulate mimicry of social but not nonsocial rewards. *Autism Research: Official Journal of the International Society for Autism Research*, *6*(6), 614–620. https://doi.org/10.1002/aur.1323

Hamilton, A. F. de C. (2013). Reflecting on the mirror neuron system in autism: A systematic review of current theories. *Developmental Cognitive Neuroscience*, *3*, 91–105. https://doi.org/10.1016/j.dcn.2012.09.008

Hamilton, A. F. de C. (2016). Gazing at me: The importance of social meaning in understanding direct-gaze cues. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1686), 20150080. https://doi.org/10.1098/rstb.2015.0080

Hamilton, A. F. de C., Brindley, R. M., & Frith, U. (2007). Imitation and action understanding in autistic spectrum disorders: How valid is the hypothesis of a deficit in the mirror neuron system? *Neuropsychologia*, *45*(8), 1859–1868. https://doi.org/10.1016/j.neuropsychologia.2006.11.022

Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: What can they tell us about social neuroscience, theory of mind and autism? *Culture and Brain*, *4*(2), 159–177. https://doi.org/10.1007/s40167-016-0044-5

Happé, F., Cook, J. L., & Bird, G. (2017). The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. *Annual Review of Psychology*, *68*(1), 243–267. https://doi.org/10.1146/annurev-psych-010416-044046

Harari-Dahan, O., & Bernstein, A. (2014). A general approach-avoidance hypothesis of Oxytocin: Accounting for social and non-social effects of oxytocin. *Neuroscience & Biobehavioral Reviews*, *47*, 506–519. https://doi.org/10.1016/j.neubiorev.2014.10.007

Harms, C., & Lakens, D. (2018). Making "null effects" informative: Statistical techniques and inferential frameworks. *Journal of Clinical and Translational Research*, *3*(Suppl 2), 382–393.

Harrell, F. E., & Dupont, C. (2018). *Hmisc: Harrell Miscellaneous* (R package version 4.1-1) [Computer software]. https://CRAN.R-project.org/package=Hmisc

Harris, D. A., Hayes-Skelton, S. A., & Ciaramitaro, V. M. (2016). What's in a Face? How Face Gender and Current Affect Influence Perceived Emotion. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01468

Hess, U., Arslan, R., Mauersberger, H., Blaison, C., Dufner, M., Denissen, J. J. A., & Ziegler, M. (2017). Reliability of surface facial electromyography: Reliability of facial EMG. *Psychophysiology*, *54*(1), 12–23. https://doi.org/10.1111/psyp.12676

Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, *40*(2), 129–141. https://doi.org/10.1016/S0167-8760(00)00161-6

Hess, U., Blaison, C., & Dandeneau, S. (2017). The impact of rewards on empathic accuracy and emotional mimicry. *Motivation and Emotion*, *41*(1), 107–112. https://doi.org/10.1007/s11031-016-9590-6

Hess, U., & Bourgeois, P. (2010). You smile–I smile: Emotion expression in social interaction. *Biological Psychology*, *84*(3), 514–520. https://doi.org/10.1016/j.biopsycho.2009.11.001

Hess, U., & Fischer, A. (2013). Emotional Mimicry as Social Regulation. *Personality and Social Psychology Review*, *17*(2), 142–157. https://doi.org/10.1177/1088868312472607

Hess, U., & Fischer, A. (2014). Emotional Mimicry: Why and When We Mimic Emotions. *Social and Personality Psychology Compass*, *8*(2), 45–57. https://doi.org/10.1111/spc3.12083

Hessels, R. S., Holleman, G. A., Cornelissen, T. H. W., Hooge, I. T. C., & Kemner, C. (2018). Eye contact takes two – autistic and social anxiety traits predict gaze behavior in dyadic interaction. *Journal of Experimental Psychopathology*, *9*(2), jep.062917. https://doi.org/10.5127/jep.062917

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, *137*(3), 463–483. https://doi.org/10.1037/a0022288

Heynen, E. J. E., Helm, G. H. P. V. der, Stams, G. J. J. M., & Korebrits, A. M. (2016). Measuring Empathy in a German Youth Prison: A Validation of the German Version of the Basic Empathy Scale (BES) in a Sample of Incarcerated Juvenile Offenders. *Journal of Forensic Psychology Practice*, *16*(5), 336–346. https://doi.org/10.1080/15228932.2016.1219217

Hoffmann, F., Banzhaf, C., Kanske, P., Gärtner, M., Bermpohl, F., & Singer, T. (2016). Empathy in depression: Egocentric and altercentric biases and the role of alexithymia. *Journal of Affective Disorders*, *199*, 23–29. https://doi.org/10.1016/j.jad.2016.03.007

Hoffmann, F., Koehne, S., Steinbeis, N., Dziobek, I., & Singer, T. (2016). Preserved Self-other Distinction During Empathy in Autism is Linked to Network Integrity of Right Supramarginal Gyrus. *Journal of Autism and Developmental Disorders*, *46*(2), 637–648. https://doi.org/10.1007/s10803-015-2609-0

Hoffmann, F., Singer, T., & Steinbeis, N. (2015). Children's Increased Emotional Egocentricity Compared to Adults Is Mediated by Age-Related Differences in Conflict Processing. *Child Development*, *86*(3), 765–780. https://doi.org/10.1111/cdev.12338

Hogeveen, J., Obhi, S. S., Banissy, M. J., Santiesteban, I., Press, C., Catmur, C., & Bird, G. (2015). Task-dependent and distinct roles of the temporoparietal junction and inferior frontal cortex in the control of imitation. *Social Cognitive and Affective Neuroscience*, *10*(7), 1003–1009. https://doi.org/10.1093/scan/nsu148

Holland, A. C., O'Connell, G., & Dziobek, I. (2021). Facial mimicry, empathy, and emotion recognition: A meta-analysis of correlations. *Cognition and Emotion*, *35*(1), 150–168. https://doi.org/10.1080/02699931.2020.1815655

Hope, R. M. (2013). *Rmisc: Ryan Miscellaneous*. https://CRAN.R-project.org/package=Rmisc

Hudry, K., & Slaughter, V. (2009). Agent familiarity and emotional context influence the everyday empathic responding of young children with autism. *Research in Autism Spectrum Disorders*, *3*(1), 74–85. https://doi.org/10.1016/j.rasd.2008.04.004

Hühnel, I., Kuszynski, J., Asendorpf, J. B., & Hess, U. (2018). Emotional mimicry of older adults' expressions: Effects of partial inclusion in a Cyberball paradigm. *Cognition and Emotion*, *32*(1), 92–101. https://doi.org/10.1080/02699931.2017.1284046

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., Dziobek, I., Gallinat, J., Wagner, M., Maier, W., & Kendrick, K. M. (2010). Oxytocin Enhances Amygdala-Dependent, Socially Reinforced Learning and Emotional Empathy in Humans. *Journal of Neuroscience*, *30*(14), 4999–5007. https://doi.org/10.1523/JNEUROSCI.5538-09.2010

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, *7*(12), 942–951. https://doi.org/10.1038/nrn2024

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Ishihara, S. (1972). *Tests for Colour-Blindness*. Kanehara Shuppan Co.Ltd.

Israelashvili, J., Sauter, D., & Fischer, A. (2019). How Well Can We Assess Our Ability to Understand Others' Feelings? Beliefs About Taking Others' Perspectives and Actual Understanding of Others' Emotions. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02475

Jackson, M. C., & Arlegui-Prieto, M. (2016). Variation in normal mood state influences sensitivity to dynamic changes in emotional expression. *Emotion*, *16*(2), 145–149. https://doi.org/10.1037/emo0000126

JASP Team. (2020). *JASP (Version 0.14.1)*.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*(5), 524–532. https://doi.org/10.1177/0956797611430953

Joober, R., Schmitz, N., Annable, L., & Boksa, P. (2012). Publication bias: What are the challenges and can they be overcome? *Journal of Psychiatry & Neuroscience*, *37*(3), 149–152. https://doi.org/10.1503/jpn.120065

Juckel, G., Heinisch, C., Welpinghus, A., & Brüne, M. (2018). Understanding Another Person's Emotions—An Interdisciplinary Research Approach. *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00414

Kampis, D., & Southgate, V. (2020). Altercentric Cognition: How Others Influence Our Cognitive Processing. *Trends in Cognitive Sciences*, *24*(11), 945–959. https://doi.org/10.1016/j.tics.2020.09.003

Kassambara, A. (2020). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* (R package version 0.6.0) [Computer software]. https://CRAN.R-project.org/package=rstatix

Kavanagh, L. C., & Winkielman, P. (2016). The Functionality of Spontaneous Mimicry and Its Influences on Affiliation: An Implicit Socialization Account. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00458

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25–41. https://doi.org/10.1016/S0010-0277(03)00064-7

Kleberg, J. L., Högström, J., Nord, M., Bölte, S., Serlachius, E., & Falck-Ytter, T. (2017). Autistic Traits and Symptoms of Social Anxiety are Differentially Related to Attention to Others' Eyes in Social Anxiety Disorder. *Journal of Autism and Developmental Disorders*, *47*(12), 3814–3821. https://doi.org/10.1007/s10803-016-2978-z

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., … Nosek, B. A. (2018). Many Labs 2: Investigating Variation in Replicability Across Samples and Settings. *Advances in Methods and Practices in Psychological Science*, *1*(4), 443–490. https://doi.org/10.1177/2515245918810225

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental Effects and Individual Differences in Linear Mixed Models: Estimating the Relationship between Spatial, Object, and Attraction Effects in Visual Attention. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00238

Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical reflexive gaze patterns on emotional faces in autism spectrum disorders. *The Journal of*

Neuroscience: The Official Journal of the Society for Neuroscience, *30*(37), 12281–12287. https://doi.org/10.1523/JNEUROSCI.0688-10.2010

Kliemann, D., Rosenblau, G., Boelte, S., Heekeren, H. R., & Dziobek, I. (2013). Face puzzle—Two new video-based tasks for measuring explicit and implicit aspects of facial emotion recognition. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00376

Korb, S., Goldman, R., Davidson, R. J., & Niedenthal, P. M. (2019). Increased Medial Prefrontal Cortex and Decreased Zygomaticus Activation in Response to Disliked Smiles Suggest Top-Down Inhibition of Facial Mimicry. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01715

Korb, S., Malsert, J., Strathearn, L., Vuilleumier, P., & Niedenthal, P. (2016). Sniff and mimic—Intranasal oxytocin increases facial mimicry in a sample of men. *Hormones and Behavior*, *84*, 64–74. https://doi.org/10.1016/j.yhbeh.2016.06.003

Kraaijenvanger, E. J., Hofman, D., & Bos, P. A. (2017). A neuroendocrine account of facial mimicry and its dynamic modulation. *Neuroscience & Biobehavioral Reviews*, *77*, 98–106. https://doi.org/10.1016/j.neubiorev.2017.03.006

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS). [Investigations with a German version of the Positive and Negative Affect Schedule (PANAS).]. *Diagnostica*, *42*(2), 139–156.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial: *Advances in Methods and Practices in Psychological Science*. https://doi.org/10.1177/2515245918770963

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, *14*(4), 334–339. https://doi.org/10.1111/1467-9280.14481

Lakin, J. L., Chartrand, T. L., & Arkin, R. M. (2008). I am too just like you: Nonconscious mimicry as an automatic behavioral response to social exclusion. *Psychological Science*, *19*(8), 816–822. https://doi.org/10.1111/j.1467-9280.2008.02162.x

Lamm, C., Bukowski, H., & Silani, G. (2016). From shared to distinct self–other representations in empathy: Evidence from neurotypical function and socio-cognitive disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150083. https://doi.org/10.1098/rstb.2015.0083

Lane, A., Luminet, O., Nave, G., & Mikolajczak, M. (2016). Is there a Publication Bias in Behavioural Intranasal Oxytocin Research on Humans? Opening the File Drawer of One Laboratory. *Journal of Neuroendocrinology*, *28*(4). https://doi.org/10.1111/jne.12384

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. NIMH, Center for the Study of Emotion & Attention.

Lawrence, M. A. (2016). *ez: Easy Analysis and Visualization of Factorial Experiments* (R package version 4.4-0) [Computer software]. https://CRAN.R-project.org/package=ez

Lawrie, L., Jackson, M. C., & Phillips, L. H. (2019). Effects of induced sad mood on facial emotion perception in young and older adults. *Aging, Neuropsychology, and Cognition*, *26*(3), 319–335. https://doi.org/10.1080/13825585.2018.1438584

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Lee, T. M. C., Ng, E. H. H., Tang, S. W., & Chan, C. C. H. (2008). Effects of sad mood on facial emotion recognition in Chinese people. *Psychiatry Research*, *159*(1), 37–43. https://doi.org/10.1016/j.psychres.2007.04.022

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*(8), 1279–1292. https://doi.org/10.3758/BF03194543

Lehrl, S. (2005). *Siegfried Lehrl: Mehrfachwahl-Wortschatz-Intelligenztest MWT-B. 5. Unveränderte Auflage. Spitta Verlag, Balingen 2005* (unveränderte Auflage, Vol. 5). Spitta Verlag.

Leighton, J., Bird, G., Orsini, C., & Heyes, C. (2010). Social attitudes modulate automatic imitation. *Journal of Experimental Social Psychology*, *46*(6), 905–910. https://doi.org/10.1016/j.jesp.2010.07.001

Leiner, D. J. (2018). *SoSci Survey* (2.6.00) [Computer software]. https://www.soscisurvey.de

Leng, G., & Ludwig, M. (2016). Intranasal Oxytocin: Myths and Delusions. *Biological Psychiatry*, *79*(3), 243–250. https://doi.org/10.1016/j.biopsych.2015.05.003

Lenth, R. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. https://CRAN.R-project.org/package=emmeans

Leppanen, J., Ng, K. W., Tchanturia, K., & Treasure, J. (2017). Meta-analysis of the effects of intranasal oxytocin on interpretation and expression of emotions. *Neuroscience and Biobehavioral Reviews*, *78*, 125–144. https://doi.org/10.1016/j.neubiorev.2017.04.010

Lieberman, M. D. (2007). Social Cognitive Neuroscience: A Review of Core Processes. *Annual Review of Psychology*, *58*(1), 259–289. https://doi.org/10.1146/annurev.psych.58.110405.085654

Likowski, K. U., Muehlberger, A., Gerdes, A. B. M., Wieser, M. J., Pauli, P., & Weyers, P. (2012). Facial mimicry and the mirror neuron system: Simultaneous acquisition of facial electromyography and functional magnetic resonance imaging. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00214

Likowski, K. U., Mühlberger, A., Seibt, B., Pauli, P., & Weyers, P. (2008). Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology*, *44*(4), 1065–1072. https://doi.org/10.1016/j.jesp.2007.10.007

Lindner, J. L., & Rosén, L. A. (2006). Decoding of Emotion through Facial Expression, Prosody and Verbal Content in Children and Adolescents with Asperger's Syndrome. *Journal of Autism and Developmental Disorders*, *36*(6), 769–777. https://doi.org/10.1007/s10803-006-0105-2

Liuzza, M. T., Vecchione, M., Dentale, F., Crostella, F., Barbaranelli, C., Caprara, G. V., & Aglioti, S. M. (2013). A look into the ballot box: Gaze following conveys information about implicit attitudes toward politicians. *Quarterly Journal of Experimental Psychology (2006)*, *66*(2), 209–216. https://doi.org/10.1080/17470218.2012.754909

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01171

Lüdecke, D. (2020). *sjPlot: Data Visualization for Statistics in Social Science* (R package version 2.8.4) [Computer software]. https://CRAN.R-project.org/package=sjPlot

Lundström, S. (2012). Autism Spectrum Disorders and Autisticlike Traits: Similar Etiology in the Extreme End and the Normal Variation. *Archives of General Psychiatry*, *69*(1), 46. https://doi.org/10.1001/archgenpsychiatry.2011.144

Lynn, S. K., Hoge, E. A., Fischer, L. E., Barrett, L. F., & Simon, N. M. (2014). Gender differences in oxytocin-associated disruption of decision bias during emotion perception. *Psychiatry Research*, *219*(1), 198–203. https://doi.org/10.1016/j.psychres.2014.04.031

Ma, Y., Shamay-Tsoory, S., Han, S., & Zink, C. F. (2016). Oxytocin and Social Adaptation: Insights from Neuroimaging Studies of Healthy and Clinical Populations. *Trends in Cognitive Sciences*, *20*(2), 133–145. https://doi.org/10.1016/j.tics.2015.10.009

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2013). *The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database.*

Marneweck, M., Loftus, A., & Hammond, G. (2013). Psychophysical Measures of Sensitivity to Facial Expression of Emotion. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00063

Marsh, L. E., Bird, G., & Catmur, C. (2016). The imitation game: Effects of social cues on "imitation" are domain-general in nature. *NeuroImage*, *139*, 368–375. https://doi.org/10.1016/j.neuroimage.2016.06.050

Martin, A. K., Huang, J., Hunold, A., & Meinzer, M. (2018). Dissociable Roles Within the Social Brain for Self–Other Processing: A HD-tDCS Study. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhy238

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Matt, G. E., Vázquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, *12*(2), 227–255. https://doi.org/10.1016/0272-7358(92)90116-P

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, *36*(4), 455–470.

McIntosh, D. N. (2006). Spontaneous facial mimicry, liking and emotional contagion. *Polish Psychological Bulletin*, *37*(1), 31–42.

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., & Wilbarger, J. L. (2006). When the social mirror breaks: Deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Developmental Science*, *9*(3), 295–302. https://doi.org/10.1111/j.1467-7687.2006.00492.x

Melloni, M., Lopez, V., & Ibanez, A. (2014). Empathy and contextual social cognition. *Cognitive, Affective & Behavioral Neuroscience*, *14*(1), 407–425. https://doi.org/10.3758/s13415-013-0205-3

Meltzoff, A. N. (2007). 'Like me': A foundation for social cognition. *Developmental Science*, *10*(1), 126–134. https://doi.org/10.1111/j.1467-7687.2007.00574.x

Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: Social neuropeptides for translational medicine. *Nature Reviews Neuroscience*, *12*(9), 524–538. https://doi.org/10.1038/nrn3044

Mier, D., Lis, S., Neuthe, K., Sauer, C., Esslinger, C., Gallhofer, B., & Kirsch, P. (2010). The involvement of emotion recognition in affective theory of mind. *Psychophysiology*, *47*(6), 1028–1039. https://doi.org/10.1111/j.1469-8986.2010.01031.x

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1309–1316. https://doi.org/10.1098/rstb.2008.0318

Montalan, B., Lelard, T., Godefroy, O., & Mouras, H. (2012). Behavioral Investigation of the Influence of Social Categorization on Empathy for Pain: A Minimal Group Paradigm Study. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00389

Montoya, R. M., Horton, R. S., Vevea, J. L., Citkowicz, M., & Lauber, E. A. (2017). A re-examination of the mere exposure effect: The influence of repeated exposure on

recognition, familiarity, and liking. *Psychological Bulletin*, *143*(5), 459–498. https://doi.org/10.1037/bul0000085

Murata, A., Saito, H., Schug, J., Ogawa, K., & Kameda, T. (2016). Spontaneous Facial Mimicry Is Enhanced by the Goal of Inferring Emotional States: Evidence for Moderation of "Automatic" Mimicry by Higher Cognitive Processes. *PLOS ONE*, *11*(4), e0153128. https://doi.org/10.1371/journal.pone.0153128

Murray, A. L., Booth, T., McKenzie, K., Kuenssberg, R., & O'Donnell, M. (2014). Are Autistic Traits Measured Equivalently in Individuals With and Without An Autism Spectrum Disorder? An Invariance Analysis of the Autism Spectrum Quotient Short Form. *Journal of Autism and Developmental Disorders*, *44*(1), 55–64. https://doi.org/10.1007/s10803-013-1851-6

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*(134), 20170213. https://doi.org/10.1098/rsif.2017.0213

Nave, G., Camerer, C., & McCullough, M. (2015). Does Oxytocin Increase Trust in Humans? A Critical Review of Research. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *10*(6), 772–789. https://doi.org/10.1177/1745691615600138

Neufeld, J., & Chakrabarti, B. (2016). Empathy Modulates the Rewarding Effect of Mimicry. *Scientific Reports*, *6*(1), 27751. https://doi.org/10.1038/srep27751

Niedenthal, P. M. (2007). Embodying Emotion. *Science*, *316*(5827), 1002–1005. https://doi.org/10.1126/science.1136930

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, Å. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition & Emotion*, *15*(6), 853–864. https://doi.org/10.1080/02699930143000194

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, Å. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology*, *30*(2), 211–222. https://doi.org/10.1002/(SICI)1099-0992(200003/04)30:2<211::AID-EJSP988>3.0.CO;2-3

Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and Brain Sciences*, *33*(6), 417–433. https://doi.org/10.1017/S0140525X10000865

Nieuwenhuis, R., Grotenhuis, H. F. te, & Pelzer, B. J. (2012). influence.ME: Tools for detecting influential data in mixed effects models. *47*. https://repository.ubn.ru.nl/handle/2066/103101

Noah, T., Schul, Y., & Mayo, R. (2018). When both the original study and its failed replication are correct: Feeling observed eliminates the facial-feedback effect. *Journal of Personality and Social Psychology*, *114*(5), 657–664. https://doi.org/10.1037/pspa0000121

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin*, *133*(2), 310–327. https://doi.org/10.1037/0033-2909.133.2.310

Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: The mirror neuron hypothesis. *Neuropsychologia*, *46*(5), 1558–1565. https://doi.org/10.1016/j.neuropsychologia.2008.01.010

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience*, *2*(3–4), 167–178. https://doi.org/10.1080/17470910701391943

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2009). Slow echo: Facial EMG evidence for the delay of spontaneous, but not voluntary, emotional mimicry in children with autism spectrum disorders. *Developmental Science*, *12*(4), 510–520. https://doi.org/10.1111/j.1467-7687.2008.00796.x

O'Brien, E., & Ellsworth, P. C. (2012). More Than Skin Deep: Visceral States Are Not Projected Onto Dissimilar Others. *Psychological Science*, *23*(4), 391–396. https://doi.org/10.1177/0956797611432179

Ohlsen, G., Zoest, W. van, & Vugt, M. van. (2013). Gender and Facial Dominance in Gaze Cuing: Emotional Context Matters in the Eyes That We Follow. *PLOS ONE*, *8*(4), e59471. https://doi.org/10.1371/journal.pone.0059471

Olszanowski, M., Wróbel, M., & Hess, U. (2020). Mimicking and sharing emotions: A re-examination of the link between facial mimicry and emotional contagion. *Cognition and Emotion*, *34*(2), 367–376. https://doi.org/10.1080/02699931.2019.1611543

Paulus, C. (2009). *Der Saarbrücker Persönlichkeitsfragebogen SPF(IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index*. http://nbn-resolving.de/urn:nbn:de:bsz:291-psydok-23630

Pavarini, G., Sun, R., Mahmoud, M., Cross, I., Schnall, S., Fischer, A., Deakin, J., Ziauddeen, H., Kogan, A., & Vuillier, L. (2019). The Role of Oxytocin in the Facial Mimicry of Affiliative vs. Non-Affiliative Emotions. *Psychoneuroendocrinology*, 104377. https://doi.org/10.1016/j.psyneuen.2019.104377

Pearson, A., Ropar, D., & Hamilton, A. F. de C. (2013). A review of visual perspective taking in autism spectrum disorder. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00652

Pedersen, T. L. (2019). *patchwork: The Composer of Plots* (R package version 1.0.0) [Computer software]. https://CRAN.R-project.org/package=patchwork

Press, C., Richardson, D., & Bird, G. (2010). Intact imitation of emotional facial actions in autism spectrum conditions. *Neuropsychologia*, *48*(11), 3291–3297. https://doi.org/10.1016/j.neuropsychologia.2010.07.012

Prins, N., & Kingdom, F. A. A. (2018). Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01250

Prinsen, J., & Alaerts, K. (2019). Eye contact enhances interpersonal motor resonance: Comparing video stimuli to a live two-person action context. *Social Cognitive and Affective Neuroscience*, *14*(9), 967–976. https://doi.org/10.1093/scan/nsz064

Prinsen, J., Bernaerts, S., Wang, Y., de Beukelaar, T. T., Cuypers, K., Swinnen, S. P., & Alaerts, K. (2017). Direct eye contact enhances mirroring of others' movements: A transcranial magnetic stimulation study. *Neuropsychologia*, *95*, 111–118. https://doi.org/10.1016/j.neuropsychologia.2016.12.011

Qiao-Tasserit, E., Quesada, M. G., Antico, L., Bavelier, D., Vuilleumier, P., & Pichon, S. (2017). Transient emotional events and individual affective traits affect emotion recognition in a perceptual decision-making task. *PLoS ONE*, *12*(2). https://doi.org/10.1371/journal.pone.0171375

Quesque, F., & Brass, M. (2019). The Role of the Temporoparietal Junction in Self-Other Distinction. *Brain Topography*, *32*(6), 943–955. https://doi.org/10.1007/s10548-019-00737-5

Quesque, F., & Rossetti, Y. (2020). What Do Theory-of-Mind Tasks Actually Measure? Theory and Practice. *Perspectives on Psychological Science*, *15*(2), 384–396. https://doi.org/10.1177/1745691619896607

Quintana, D. S. (2018). Revisiting non-significant effects of intranasal oxytocin using equivalence testing. *Psychoneuroendocrinology*, *87*, 127–130. https://doi.org/10.1016/j.psyneuen.2017.10.010

Quintana, D. S., Alvares, G. A., Hickie, I. B., & Guastella, A. J. (2015). Do delivery routes of intranasally administered oxytocin account for observed effects on social cognition and behavior? A two-level model. *Neuroscience and Biobehavioral Reviews*, *49*, 182–192. https://doi.org/10.1016/j.neubiorev.2014.12.011

Quintana, D. S., & Guastella, A. J. (2020). An Allostatic Theory of Oxytocin. *Trends in Cognitive Sciences*, *0*(0). https://doi.org/10.1016/j.tics.2020.03.008

Quintana, D. S., Lischke, A., Grace, S., Scheele, D., Ma, Y., & Becker, B. (2021). Advances in the field of intranasal oxytocin research: Lessons learned and future directions for clinical research. *Molecular Psychiatry*, *26*(1), 80–91. https://doi.org/10.1038/s41380-020-00864-7

Quintana, D. S., Smerud, K. T., Andreassen, O. A., & Djupesland, P. G. (2018). Evidence for intranasal oxytocin delivery to the brain: Recent advances and future perspectives. *Therapeutic Delivery*, *9*(7), 515–525. https://doi.org/10.4155/tde-2018-0002

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rauchbauer, B., Majdandžić, J., Stieger, S., & Lamm, C. (2016). The Modulation of Mimicry by Ethnic Group-Membership and Emotional Expressions. *PLOS ONE*, *11*(8), e0161064. https://doi.org/10.1371/journal.pone.0161064

Ready, R. E., Clark, L. A., Watson, D., & Westerhouse, K. (2000). Self- and peer-related personality: Agreement, trait ratability, and the "self-based heuristic." *Journal of Research in Personality*, *34*(2), 208–224. https://doi.org/10.1006/jrpe.1999.2280

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience*, *20*(8), 495–505. https://doi.org/10.1038/s41583-019-0179-4

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Riegel, M., Żurawski, Ł., Wierzba, M., Moslehi, A., Klocek, Ł., Horvat, M., Grabowska, A., Michałowski, J., Jednoróg, K., & Marchewka, A. (2016). Characterization of the Nencki Affective Picture System by discrete emotional categories (NAPS BE). *Behavior Research Methods*, *48*(2), 600–612. https://doi.org/10.3758/s13428-015-0620-1

Rilling, J. K., DeMarco, A. C., Hackett, P. D., Chen, X., Gautam, P., Stair, S., Haroon, E., Thompson, R., Ditzen, B., Patel, R., & Pagnoni, G. (2014). Sex differences in the neural and behavioral response to intranasal oxytocin and vasopressin during human social interaction. *Psychoneuroendocrinology*, *39*, 237–248. https://doi.org/10.1016/j.psyneuen.2013.09.022

Riva, F., Triscoli, C., Lamm, C., Carnaghi, A., & Silani, G. (2016). Emotional Egocentricity Bias Across the Life-Span. *Frontiers in Aging Neuroscience*, *8*. https://doi.org/10.3389/fnagi.2016.00074

Robbins, J. M., & Krueger, J. I. (2005). Social Projection to Ingroups and Outgroups: A Review and Meta-Analysis. *Personality and Social Psychology Review*, *9*(1), 32–47. https://doi.org/10.1207/s15327957pspr0901_3

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happé, F., Plomin, R., & Ronald, A. (2011). Evidence That Autistic Traits Show the Same Etiology in the General Population and at the Quantitative Extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry*, *68*(11), 1113–1121. https://doi.org/10.1001/archgenpsychiatry.2011.119

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301. https://doi.org/10.1016/0022-1031(77)90049-X

RStudio Team. (2019). *RStudio: Integrated Development Environment for R*. RStudio, Inc. http://www.rstudio.com/

Rymarczyk, K., Biele, C., Grabowska, A., & Majczynski, H. (2011). EMG activity in response to static and dynamic facial expressions. *International Journal of Psychophysiology*, *79*(2), 330–333. https://doi.org/10.1016/j.ijpsycho.2010.11.001

Ryskin, R. A., & Brown-Schmidt, S. (2014). Do Adults Show a Curse of Knowledge in False-Belief Reasoning? A Robust Estimate of the True Effect Size. *PLoS ONE*, *9*(3), e92406. https://doi.org/10.1371/journal.pone.0092406

Sali, A. W., Anderson, B. A., & Yantis, S. (2014). The Role of Reward Prediction in the Control of Attention. *Journal of Experimental Psychology. Human Perception and Performance*, *40*(4), 1654–1664. https://doi.org/10.1037/a0037267

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255–1266. https://doi.org/10.1037/a0018729

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2012). Enhancing Social Ability by Stimulating Right Temporoparietal Junction. *Current Biology*, *22*(23), 2274–2277. https://doi.org/10.1016/j.cub.2012.10.018

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, *122*(2), 228–235. https://doi.org/10.1016/j.cognition.2011.11.004

Sato, W., & Yoshikawa, S. (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, *104*(1), 1–18. https://doi.org/10.1016/j.cognition.2006.05.001

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(4), 393–414. https://doi.org/10.1017/S0140525X12000660

Schmid, P. C., & Schmid Mast, M. (2010). Mood effects on emotion recognition. *Motivation and Emotion*, *34*(3), 288–292. https://doi.org/10.1007/s11031-010-9170-0

Schmidt, S. (2009). Shall we Really do it Again? The Powerful Concept of Replication is Neglected in the Social Sciences. *Review of General Psychology*, *13*(2), 90–100. https://doi.org/10.1037/a0015108

Schrammel, F., Pannasch, S., Graupner, S.-T., Mojzisch, A., & Velichkovsky, B. M. (2009). Virtual friend or threat? The effects of facial expression and gaze interaction on

psychophysiological responses and emotional experience. *Psychophysiology*, *46*(5), 922–931. https://doi.org/10.1111/j.1469-8986.2009.00831.x

Schulte-Rüther, M., Otte, E., Adigüzel, K., Firk, C., Herpertz-Dahlmann, B., Koch, I., & Konrad, K. (2017). Intact mirror mechanisms for automatic facial emotions in children and adolescents with autism spectrum disorder: Intact mirror mechanisms in Austism. *Autism Research*, *10*(2), 298–310. https://doi.org/10.1002/aur.1654

Schulze, L., Renneberg, B., & Lobmaier, J. S. (2013). Gaze perception in social anxiety and social anxiety disorder. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00872

Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion recognition in faces. *Journal of Vision*, *14*(13), 14–14. https://doi.org/10.1167/14.13.14

Seibt, B., Mühlberger, A., Likowski, K. U., & Weyers, P. (2015). Facial mimicry in its social setting. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01122

Shahrestani, S., Kemp, A. H., & Guastella, A. J. (2013). The Impact of a Single Administration of Intranasal Oxytocin on the Recognition of Basic Emotions in Humans: A Meta-Analysis. *Neuropsychopharmacology*, *38*(10), 1929–1936. https://doi.org/10.1038/npp.2013.86

Shamay-Tsoory, S. G., & Abu-Akel, A. (2016). The Social Salience Hypothesis of Oxytocin. *Biological Psychiatry*, *79*(3), 194–202. https://doi.org/10.1016/j.biopsych.2015.07.020

Siedlecka, E., & Denson, T. F. (2019). Experimental Methods for Inducing Basic Emotions: A Qualitative Review. *Emotion Review*, *11*(1), 87–97. https://doi.org/10.1177/1754073917749016

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: An fMRI study. *Social Neuroscience*, *3*(2), 97–112. https://doi.org/10.1080/17470910701577020

Silani, G., Lamm, C., Ruff, C. C., & Singer, T. (2013). Right Supramarginal Gyrus Is Crucial to Overcome Emotional Egocentricity Bias in Social Judgments. *Journal of Neuroscience*, *33*(39), 15466–15476. https://doi.org/10.1523/JNEUROSCI.1488-13.2013

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation. *Cognition*, *166*, 371–381. https://doi.org/10.1016/j.cognition.2017.06.003

Sims, T. B., Neufeld, J., Johnstone, T., & Chakrabarti, B. (2014). Autistic traits modulate frontostriatal connectivity during processing of rewarding faces. *Social Cognitive and Affective Neuroscience*, *9*(12), 2010–2016. https://doi.org/10.1093/scan/nsu010

Sims, T. B., Van Reekum, C. M., Johnstone, T., & Chakrabarti, B. (2012). How reward modulates mimicry: EMG evidence of greater facial mimicry of more rewarding happy faces:

Greater mimicry of rewarding faces. *Psychophysiology*, *49*(7), 998–1004. https://doi.org/10.1111/j.1469-8986.2012.01377.x

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2019). *afex: Analysis of Factorial Experiments.* (R package version 0.23-0) [Computer software]. https://CRAN.R-project.org/package=afex

Smith, E. R., & Semin, G. R. (2007). Situated Social Cognition. *Current Directions in Psychological Science*, *16*(3), 132–135. https://doi.org/10.1111/j.1467-8721.2007.00490.x

Sonnby–Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, *43*(5), 433–443. https://doi.org/10.1111/1467-9450.00312

Sowden, S., & Catmur, C. (2015). The role of the right temporoparietal junction in the control of imitation. *Cerebral Cortex*, *25*(4), 1107–1113. https://doi.org/10.1093/cercor/bht306

Sowden, S., Koehne, S., Catmur, C., Dziobek, I., & Bird, G. (2016). Intact Automatic Imitation and Typical Spatial Compatibility in Autism Spectrum Disorder: Challenging the Broken Mirror Theory. *Autism Research: Official Journal of the International Society for Autism Research*, *9*(2), 292–300. https://doi.org/10.1002/aur.1511

Sowden, S., & Shah, P. (2014). Self-other control: A candidate mechanism for social cognitive function. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00789

Spain, D., Sin, J., Linder, K. B., McMahon, J., & Happé, F. (2018). Social anxiety in autism spectrum disorder: A systematic review. *Research in Autism Spectrum Disorders*, *52*, 51–68. https://doi.org/10.1016/j.rasd.2018.04.007

Spengler, F. B., Schultz, J., Scheele, D., Essel, M., Maier, W., Heinrichs, M., & Hurlemann, R. (2017). Kinetics and Dose Dependency of Intranasal Oxytocin Effects on Amygdala Reactivity. *Biological Psychiatry*, *82*(12), 885–894. https://doi.org/10.1016/j.biopsych.2017.04.015

Spengler, S., Bird, G., & Brass, M. (2010). Hyperimitation of Actions Is Related to Reduced Understanding of Others' Minds in Autism Spectrum Conditions. *Biological Psychiatry*, *68*(12), 1148–1155. https://doi.org/10.1016/j.biopsych.2010.09.017

Spengler, S., Cramon, D. Y. von, & Brass, M. (2010). Resisting motor mimicry: Control of imitation involves processes central to social cognition in patients with frontal and temporo-parietal lesions. *Social Neuroscience*, *5*(4), 401–416. https://doi.org/10.1080/17470911003687905

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Control of shared representations relies on key processes involved in mental state attribution. *Human Brain Mapping*, *30*(11), 3704–3718. https://doi.org/10.1002/hbm.20800

Stangier, U., Heidenreich, T., Berardi, A., Golbs, U., & Hoyer, J. (1999). Die Erfassung sozialer Phobie durch Social Interaction Anxiety Scale (SIAS) und die Social Phobia Scale (SPS).

[Assessment of social phobia by the Social Interaction Anxiety Scale (SIAS) and the Social Phobia Scale (SPS).]. *Zeitschrift Für Klinische Psychologie*, *28*(1), 28–36. https://doi.org/10.1026//0084-5345.28.1.28

Steinbeis, N. (2016). The role of self–other distinction in understanding others' mental and emotional states: Neurocognitive mechanisms in children and adults. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150074. https://doi.org/10.1098/rstb.2015.0074

Steinbeis, N., & Singer, T. (2014). Projecting my envy onto you: Neurocognitive mechanisms of an offline emotional egocentricity bias. *NeuroImage*, *102*, 370–380. https://doi.org/10.1016/j.neuroimage.2014.08.007

Stel, M., Blascovich, J., McCall, C., Mastop, J., van Baaren, R. B., & Vonk, R. (2009). Mimicking disliked others: Effects of a priori liking on the mimicry-liking link. *European Journal of Social Psychology*, n/a-n/a. https://doi.org/10.1002/ejsp.655

Stel, M., & van Knippenberg, A. (2008). The Role of Facial Mimicry in the Recognition of Affect. *Psychological Science*, *19*(10), 984–985. https://doi.org/10.1111/j.1467-9280.2008.02188.x

Stroebe, W., & Strack, F. (2014). The Alleged Crisis and the Illusion of Exact Replication. *Perspectives on Psychological Science*, *9*(1), 59–71. https://doi.org/10.1177/1745691613514450

Tabak, B. A., Teed, A. R., Castle, E., Dutcher, J. M., Meyer, M. L., Bryan, R., Irwin, M. R., Lieberman, M. D., & Eisenberger, N. I. (2019). Null results of oxytocin and vasopressin administration across a range of social cognitive and behavioral paradigms: Evidence from a randomized controlled trial. *Psychoneuroendocrinology*, *107*, 124–132. https://doi.org/10.1016/j.psyneuen.2019.04.019

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General*, *142*(1), 151–162. https://doi.org/10.1037/a0028232

Todd, A. R., Simpson, A. J., & Tamir, D. I. (2016). Active perspective taking induces flexible use of self-knowledge during social inference. *Journal of Experimental Psychology: General*, *145*(12), 1583–1588. https://doi.org/10.1037/xge0000237

Tomova, L., Heinrichs, M., & Lamm, C. (2019). The Other and Me: Effects of oxytocin on self-other distinction. *International Journal of Psychophysiology*, *136*, 49–53. https://doi.org/10.1016/j.ijpsycho.2018.03.008

Tomova, L., von Dawans, B., Heinrichs, M., Silani, G., & Lamm, C. (2014). Is stress affecting our ability to tune into others? Evidence for gender differences in the effects of stress on self-other distinction. *Psychoneuroendocrinology*, *43*, 95–104. https://doi.org/10.1016/j.psyneuen.2014.02.006

Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion (Washington, D.C.)*, *8*(1), 81–95. https://doi.org/10.1037/1528-3542.8.1.81

Trafimow, D. (2019). Why successful replications across contexts and Operationalizations might not be good for theory building or testing. *Journal for the Theory of Social Behaviour*, *49*(3), 359–368. https://doi.org/10.1111/jtsb.12211

Trilla Gros, I., Panasiti, M. S., & Chakrabarti, B. (2015). The plasticity of the mirror system: How reward learning modulates cortical motor simulation of others. *Neuropsychologia*, *70*, 255–262. https://doi.org/10.1016/j.neuropsychologia.2015.02.033

Trilla, I., Weigand, A., & Dziobek, I. (2020). Affective states influence emotion perception: Evidence for emotional egocentricity. *Psychological Research*. https://doi.org/10.1007/s00426-020-01314-3

Uljarevic, M., & Hamilton, A. (2013). Recognition of Emotions in Autism: A Formal Meta-Analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

Valentin, V. V., & O'Doherty, J. P. (2009). Overlapping Prediction Errors in Dorsal Striatum During Instrumental Learning With Juice and Money Reward in the Human Brain. *Journal of Neurophysiology*, *102*(6), 3384–3391. https://doi.org/10.1152/jn.91195.2008

Valstad, M., Alvares, G. A., Andreassen, O. A., Westlye, L. T., & Quintana, D. S. (2016). The relationship between central and peripheral oxytocin concentrations: A systematic review and meta-analysis protocol. *Systematic Reviews*, *5*(1), 49. https://doi.org/10.1186/s13643-016-0225-5

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(23), 6454–6459. https://doi.org/10.1073/pnas.1521897113

Van Boven, L., & Loewenstein, G. (2003). Social Projection of Transient Drive States. *Personality and Social Psychology Bulletin*, *29*(9), 1159–1168. https://doi.org/10.1177/0146167203254597

van Boxtel, A. (2010). Facial EMG as a tool for inferring affective states. In A.J. Spink, F. Grieco, O. Krips, L. Loijens, L. Noldus, & P. Zimmerman (Eds.), *Proceedings of Measuring Behavior 2010* (pp. 104–108). Noldus Information technology.

van der Schalk, J., Fischer, A., Doosje, B., Wigboldus, D., Hawk, S., Rotteveel, M., & Hess, U. (2011). Convergent and divergent responses to emotional displays of ingroup and outgroup. *Emotion*, *11*(2), 286–298. https://doi.org/10.1037/a0022582

van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion (Washington, D.C.)*, *11*(4), 907–920. https://doi.org/10.1037/a0023853

van Dillen, L. F., Harris, L. T., van Dijk, W. W., & Rotteveel, M. (2015). Looking with different eyes: The psychological meaning of categorisation goals moderates facial reactivity to facial expressions. *Cognition and Emotion*, *29*(8), 1382–1400. https://doi.org/10.1080/02699931.2014.982514

Vivanti, G., & Dissanayake, C. (2014). Propensity to Imitate in Autism Is Not Modulated by the Model's Gaze Direction: An Eye-Tracking Study. *Autism Research*, *7*(3), 392–399. https://doi.org/10.1002/aur.1376

von Mohr, M., Finotti, G., Ambroziak, K. B., & Tsakiris, M. (2019). Do you hear what I see? An audio-visual paradigm to assess emotional egocentricity bias. *Cognition & Emotion*, 1–15. https://doi.org/10.1080/02699931.2019.1683516

Walum, H., Waldman, I. D., & Young, L. J. (2016). Statistical and Methodological Considerations for the Interpretation of Intranasal Oxytocin Studies. *Biological Psychiatry*, *79*(3), 251–257. https://doi.org/10.1016/j.biopsych.2015.06.016

Wang, Y., & Hamilton, A. F. de C. (2012). Social top-down response modulation (STORM): A model of the control of mimicry in social interaction. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00153

Wang, Y., & Hamilton, A. F. de C. (2014). Why does gaze enhance mimicry? Placing gaze-mimicry effects in relation to other gaze phenomena. *Quarterly Journal of Experimental Psychology (2006)*, *67*(4), 747–762. https://doi.org/10.1080/17470218.2013.828316

Wang, Y., & Hamilton, A. F. de C. (2015). Anterior medial prefrontal cortex implements social priming of mimicry. *Social Cognitive and Affective Neuroscience*, *10*(4), 486–493. https://doi.org/10.1093/scan/nsu076

Wang, Y., Newport, R., & Hamilton, A. F. de C. (2011). Eye contact enhances mimicry of intransitive hand movements. *Biology Letters*, *7*(1), 7–10. https://doi.org/10.1098/rsbl.2010.0279

Wang, Y., Ramsey, R., & Hamilton, A. F. de C. (2011). The control of mimicry by eye contact is mediated by medial prefrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(33), 12001–12010. https://doi.org/10.1523/JNEUROSCI.0845-11.2011

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLOS ONE*, *12*(5), e0177239. https://doi.org/10.1371/journal.pone.0177239

Weigand, A., Trilla, I., Enk, L., O'Connell, G., Prehn, K., Brick, T. R., & Dziobek, I. (2021). How Much of Me Do I See in Other Minds? Modulating Egocentricity in Emotion Judgments by tDCS. *Brain Sciences*, *11*(4), 512. https://doi.org/10.3390/brainsci11040512

Weisbuch, M., Pauker, K., Adams Jr., R. B., Lamer, S. A., & Ambady, N. (2017). Race, power, and reflexive gaze following. *Social Cognition*, *35*(6), 619–638. https://doi.org/10.1521/soco.2017.35.6.619

Wickham, H. (2009). *Ggplot2*. Springer New York. https://doi.org/10.1007/978-0-387-98141-3

Willard, D. F. X., & Markman, A. B. (2017). Anchoring on Self and Others During Social Inferences. *Topics in Cognitive Science*, *9*(3), 819–841. https://doi.org/10.1111/tops.12275

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Wirth, J. H., Sacco, D. F., Hugenberg, K., & Williams, K. D. (2010). Eye gaze as relational evaluation: Averted eye gaze leads to feelings of ostracism and relational devaluation. *Personality & Social Psychology Bulletin*, *36*(7), 869–882. https://doi.org/10.1177/0146167210370032

Woo, B. M., & Mitchell, J. P. (2020). Simulation: A strategy for mindreading similar but not dissimilar others? *Journal of Experimental Social Psychology*, *90*, 104000. https://doi.org/10.1016/j.jesp.2020.104000

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger Syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders*, *35*(3), 331–335. https://doi.org/10.1007/s10803-005-3300-7

Yabar, Y., Johnston, L., Miles, L., & Peace, V. (2006). Implicit Behavioral Mimicry: Investigating the Impact of Group Membership. *Journal of Nonverbal Behavior*, *30*(3), 97–113. https://doi.org/10.1007/s10919-006-0010-6

Yik, M., Wong, K. F. E., & Zeng, K. J. (2019). Anchoring-and-Adjustment During Affect Inferences. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02567

Zhang, X., Yu, H. W., & Barrett, L. F. (2014). How does this make you feel? A comparison of four affect induction procedures. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00689

Eidesstattliche Erklärung

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben. Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad. Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde. Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin vom 5. März 2015. Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen Promotionsbearbeiterinnen/Promotions-beratern stattgefunden hat und dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

I hereby declare that I completed the doctoral thesis independently based on the stated resources and aids. I have not applied for a doctoral degree elsewhere and do not have a corresponding doctoral degree. I have not submitted the doctoral thesis, or parts of it, to another academic institution and the thesis has not been accepted or rejected. I declare that I have acknowledged the Doctoral Degree Regulations which underlie the procedure of the Faculty of Life Sciences of Humboldt-Universität zu Berlin, as amended on 5th March 2015. Furthermore, I declare that no collaboration with commercial doctoral degree supervisors took place, and that the principles of Humboldt-Universität zu Berlin for ensuring good academic practice were abided by.

Irene Trilla Gros